सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

14325

I.A.R.I.6.
GIP NLK—H-3 I.A.R.I.—10-5-55—15,000

The Faraday Society reserves the Copyright of all Communications published in the " Transactions ".

Transactions of the Faraday Society.

FOUNDED 1903.

TO PROMOTE THE STUDY OF ELECTROCHEMISTRY, ELECTROMETALLURGY, CHEMICAL PHYSICS, METALLOGRAPHY, AND KINDRED SUBJECTS.

Vol. XLII A. 1946. 20s.

A GENERAL DISCUSSION
ON
DIELECTRICS

(For Table of Contents see page 254)

GURNEY AND JACKSON
LONDON: 98 GREAT RUSSELL STREET
EDINBURGH: TWEEDDALE COURT

৬৩২২৫

OFFICERS AND COUNCIL, 1946-1947.

President.

Prof. W. E. Garner, C.B.E., D.Sc., F.R.S.

Vice-Presidents who have held the Office of President.

Prof. F. G. Donnan, C.B.E., Ph.D., F.R.S.
Prof. C. H. Desch, D.Sc., F.R.S.
Prof. N. V. Sidgwick, Sc.D., D.Sc., F.R.S.
Prof. M. W. Travers, D.Sc., F.R.S.
Prof. E. K. Rideal, M.B.E., D.Sc., F.R.S.

Vice-Presidents.

Prof. A. Findlay, D.Sc.
Sir Charles Goodeve, D.Sc., F.R.S.
R. O. Griffith, D.Sc.
Sir John Lennard-Jones, K.B.E., D.Sc., F.R.S.
R. Lessing, Ph.D.
Prof. H. W. Melville, Ph.D., F.R.S.
Prof. M. Polanyi, Ph.D., M.D., F.R.S.

Honorary Treasurer.

R. E. Slade, D.Sc.

Chairman of the Publications Committee.

Prof. A. J. Allmand, D.Sc., F.R.S.

Ordinary Members of Council.

R. P. Bell, M.A., F.R.S.
G. M. Bennett, M.A., Sc.D., F.R.S.
A. Caress, Ph.D.
Prof. C. W. Davies, D.Sc.
G. Gee, Ph.D.
G. S. Hartley, D.Sc.
F. I. G. Rawlins, M.Sc., F.R.S.E., F.Inst.P.
H. W. Thompson, M.A., D.Phil., F.R.S.
O. H. Wansbrough-Jones, O.B.E., Ph.D.
Sir Robert Robertson, K.B.E., D.Sc., F.R.S.

Honorary Secretary and Editor.

G. S. W. Marlow, B.Sc., F.R.I.C.

Assistant Editor.
F. C. Tompkins, Ph.D.

Assistant Secretary.
Miss B. E. Kornitzer.

6 Gray's Inn Square, London, W.C. 1.
(Telephone: Chancery 8101.)

Publications Committee.

Prof. A. J. Allmand, D.Sc., F.R.S. (Chairman).

Prof. N. K. Adam, M.A., Sc.D., F.R.S.
Prof. W. T. Astbury, M.A., Sc.D., F.Inst.P., F.R.S.
T. Beacall, O.B.E.
C. A. Coulson, M.A., Ph.D., D.Sc.
Prof. H. J. Emeléus, D.Sc., F.R.S.
Prof. M. G. Evans, D.Sc., F.R.S.
Prof. W. E. Garner, C.B.E., D.Sc., F.R.S.
G. Gee, Ph.D.
Prof. E. A. Guggenheim, M.A., Sc.D., F.R.S.
Prof. C. N. Hinshelwood, Sc.D., F.R.S.
Prof. W. C. M. Lewis, D.Sc., F.R.S.
Prof. H. W. Melville, Ph.D., F.R.S.
Prof. M. Polanyi, Ph.D., M.D., F.R.S.
Prof. J. T. Randall, D.Sc., F.Inst.P., F.R.S.
Prof. E. K. Rideal, M.B.E., D.Sc., F.R.S.
Sir Robert Robertson, K.B.E., D.Sc., F.R.S.
D. W. G. Style, Ph.D.
Prof. W. F. K. Wynne-Jones, D.Sc.

A GENERAL DISCUSSION

ON

DIELECTRICS

A GENERAL DISCUSSION on Dielectrics was held at the University, Bristol, on Wednesday, 24th April, from 2.15 p.m. to 6 p.m., Thursday, 25th April, from 10 a.m. to 6 p.m., and Friday, 26th April, from 10 a.m. to 1 p.m., the President, Professor W. E. Garner, F.R.S., occupying the chair throughout. About 150 members and visitors were present, and among the distinguished overseas members, the President welcomed the following :—

> Dr. P. Abadie (Bagneux), Prof. A. E. van Arkel (Leiden), Prof. E. Bauer (Paris), Dr. C. J. F. Böttcher (Leiden), Miss A. C. B. Dekking (Leiden), Prof. P. Girard (Paris), Dr. G. M. Jonker (Eindhoven), Dr. M. Magat (Paris), Monsieur D. Massignon (Paris), Dr. G. W. Nederbracht (Amsterdam), Monsieur M. Paic (Zagreb) and Dr. M. Surdin (Paris).

The papers to be discussed had all been issued in Advance Proof and were considered under the following main heads :—

A. Theoretical.
B. Theory of Relaxation Times.
C. Breakdown.
D. Measurement Technique and Terminology.
E. Applications, including Liquid Systems.
F. Long Chain Compounds.
G. Solids.

All the papers and the reports of the Discussion thereon appear in the following pages.

GENERAL INTRODUCTION

BY THE PRESIDENT

This is the third discussion of the Faraday Society to be held in this lecture theatre, the first under the chairmanship of Professor Lowry in 1929 on Molecular Spectra and Molecular Structure and the second in 1938 under the Chairmanship of Professor Travers on Chemical Reactions Involving Solids. Neither of these meetings is the true precursor of the present meeting but one held in Oxford on Dipole Moments in 1934. At that time, chemists were already awake to the importance of dielectric constants, which, together with molecular spectra and X-ray and electronic diffraction, was being actively pursued for the elucidation of the structure of molecules. The main problems left unsolved by that meeting are, strangely enough, those which form the subject of the present discussion. Thus, Debye discussed the energy absorption in dielectrics in relation to relaxation time, which is a main item on our present agenda, and there-was considerable doubt expressed as to the significance of the dipole moments as determined from measurements of liquids.

The present discussion occurs at a time when there have been great advances in technique due to developments in radio for government departments and for industry. This permits of several new lines of approach to the structure of liquids and solids and also to that state of matter which appears to be neither liquid nor crystalline, the new plastic substances.

It is one of the functions of the Faraday Society to promote like topics in the realm of physical chemistry and physics. It is rarer for engineers to be concerned in our conferences and we extend a hearty welcome to the engineers present. There is another equally important function of the Society in that it provides a means for the dispersion of new knowledge, so that it can be rapidly applied in the pure science and industrial fields.

A. THEORETICAL

DIPOLAR INTERACTION.

By H. Fröhlich.

Received 11*th March*, 1946.

1. Considerations on the interaction between dipoles usually play an important role in theoretical investigations on dipolar substances. The main difficulty in treating this interaction is its long range ($\propto 1/r^3$, r = distance between dipoles) which does not allow a restriction to the interaction between nearest neighbours. In fact the energy of a polarised substance usually depends on its shape because the number of dipoles in a distance between r and $r + dr$ from a given one increases proportional to r^2dr (isotropy assumed), *i.e.* their contribution to the interaction is proportional to dr/r. There have been many attempts to find approximate methods to calculate the influence of dipolar interaction on the dielectric properties. In some of them leading to formulæ like Onsager's or Clausius' and Mosotti's the substance is treated in a semi-macroscopic way. It will be one of the purposes of this paper to discuss the range of validity which such formulæ can claim (§ 3). A possibility to account for the interaction in a relatively simple way arises in some strongly anisotropic substances, *e.g.* in solid ketones, in which the dipoles are concentrated in certain planes (§ 5). Before discussing these questions, a survey on the properties of independent dipoles will be given.

2. Suppose that the dipolar interaction energy per dipole is small compared with kT. In this case the interaction can be neglected. This means that in the presence of an external electrical field E the field acting on a dipole is also given by E. In a liquid in the absence of an external field there is no preferential direction for a dipole. From the Boltzmann principle one then derives that the polarisation P_0 in a static external field E_0 is

$$P_0 = \frac{\mu^2 n}{3kT} E_0 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where μ = dipole moment, n = number of dipoles per c.c. In an alternating field E,

$$E = E_0 e^{i\omega t}, \quad . \quad . \quad . \quad . \quad . \quad (2)$$

(ω = angular frequency) the polarisation P is found to be

$$P = P_0 \frac{e^{i\omega t}}{1 + i\omega\tau}, \quad . \quad . \quad . \quad . \quad (3)$$

where τ is the time of relaxation. Equation (3) leads to the well-known Debye formula for dielectric loss.

In solids there are usually a discrete number of equilibrium directions for a dipole. If in the absence of an external field the energy of a dipole is the same for all these directions, then formulæ similar to (1) and (3) hold. In this case a dipole oscillates about an equilibrium position and occasionally jumps into another equilibrium position. In alternating external fields these oscillations (angular frequency ω_0) give rise to a dipolar loss which is not of the Debye type, but due to the width of the resonance absorption near ω_0. It can be shown that in this case the loss angle ϕ is given by [1]

$$\tan \phi = \frac{\Delta\epsilon}{2\epsilon_s}\left\{\frac{\omega\tau}{1 + (\omega + \omega_0)^2\tau^2} + \frac{\omega\tau}{1 + (\omega - \omega_0)^2\tau^2}\right\} \quad . \quad . \quad (4)$$

if $\Delta\epsilon \ll \epsilon_s$. Here ϵ_s is the static dielectric constant, and $\Delta\epsilon$ the contribution of the oscillating dipoles to ϵ_s. Equation (4) shows resonance near ω_0 but becomes identical with the corresponding Debye formula if $\omega \gg \omega_0$. The theory leading to equation (4) was also developed independently by Van Vleck and Weisskopf.[2]

A solid of the type discussed here must thus be expected to have two maxima for tan ϕ; a long wave one which is of the Debye type, and which is connected with the transitions of dipoles between the equilibrium directions, and a short wave maximum near the angular frequency ω_0 as described by (4). This latter is usually in the infra-red but in some cases its influence may be expected to reach into the range of cm. waves.

3. Let us now turn to a discussion of approximate methods to account for dipolar interaction. Amongst them those used by Debye (Clausius-Mosotti) and by Onsager are conspicuous for their simplicity. It seems, however, that the range within which these methods can claim validity has not always been respected, and naturally difficulties may be expected if they are applied outside this range (*e.g.* the prediction of an electric Curie point by the Debye formula).

In these methods the local field F acting on a dipole is calculated by replacing by a continuous medium the surroundings of this dipole outside a cavity with volume $1/n$. Following Onsager [3] the field F inside the cavity can be considered to consist of two parts, (i) the cavity field G which is obtained on the assumption that the dipole has been removed from the cavity, and (ii) the reaction field R which is the field due to the polarisation of the surroundings by the dipole. Thus

$$\vec{F} = \vec{G} + \vec{R} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Assuming that there is no non-dipolar polarisation, the static polarisation P_0 can be shown to be identical with that of independent dipoles in a field $E\partial F/\partial E$. Thus from (1)

$$P_0 = \frac{\mu^2 n}{3kT}\frac{\partial F}{\partial E}E. \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

For G one finds

$$G = \frac{3\epsilon}{2\epsilon + 1}E \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where ϵ is the dielectric constant. To obtain an estimate for R, assume that the dipole is fixed in a given direction for a very long time. Then the reaction field has a value R_0 given by

$$\vec{R}_0 = \frac{2(\epsilon_s - 1)}{2\epsilon_s + 1}\frac{4\pi n}{3}\vec{\mu}. \quad . \quad . \quad . \quad . \quad . \quad (8)$$

To obtain the Debye (Clausius-Mosotti) formula one has to assume that for a calculation of R, the dipole moment μ can be replaced by its average value, *i.e.* P/n. Since

$$(\epsilon - 1)E = 4\pi P, \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

one finds from (5), (7) and (8),

$$F = \frac{3\epsilon}{2\epsilon + 1}E + \frac{2(\epsilon - 1)}{2\epsilon + 1}\frac{4\pi}{3}P = \frac{\epsilon + 2}{3}E = E + \frac{4\pi}{3}P, \quad . \quad (10)$$

i.e. the Lorentz field which is the basis for Debye's procedure.[4]

---

[1] Fröhlich, *E.R.A. Report*, L/T 163, *Nature*, 1946, 157, 478.

[2] Van Vleck and Weisskopf, *Rev. Mod. Physics*, 1945, 17, 225.

[3] Onsager, *J. Amer. Chem. Soc.*, 1936, 58, 1487.

[4] Debye, *Polar Molecules*; New York, 1929.

Onsager on the other hand assumes that R has always the value R_0 and is thus independent of E. Thus from (5) and (7),

$$\vec{F} = \frac{3\epsilon}{2\epsilon + 1}\vec{E} + \vec{R}_0$$

and hence

$$\frac{\partial F}{\partial E}E = \frac{3\epsilon}{2\epsilon + 1}E = E + \frac{4\pi}{2\epsilon + 1}P. \quad . \quad . \quad . \quad (11)$$

Thus the factor $4\pi/3$ in the Lorentz field has been replaced by the smaller value $4\pi/(2\epsilon + 1)$. The dielectric constant in both cases is thus obtained from (6) and (9) using for $\partial F/\partial E$ the values obtained from (10) and (11) respectively (*i.e.* $(\epsilon + 2)/3$ and $3\epsilon/(2\epsilon + 1)$. Thus

$$\epsilon - 1 = \frac{\mu^2 4\pi n}{3kT} \times \begin{cases} \dfrac{\epsilon + 2}{3} & \text{Clausius-Mosotti} \\ \dfrac{3\epsilon}{2\epsilon + 1} & \text{Onsager} \end{cases} \quad . \quad . \quad . \quad (12)$$

Both procedures are open to objections. To justify (10) it would be necessary for a dipole to rotate so fast that the polarisation P cannot follow, *i.e.* that μ can be replaced by its average. This requires that the time of relaxation τ is large compared with the period of the rotation. While thus (10) may hold for rotating dipoles it cannot be applied to dipoles which only occasionally jump from one equilibrium direction to another. Here Onsager's treatment would seem to be more appropriate for he assumes that R can always follow the dipole. This would require that the time a dipole spends in a given direction is large compared with the relaxation time τ. On the other hand, τ is usually determined by this mean life time in one position and, therefore, Onsager's assumption cannot be true. Nevertheless, it can be shown that Onsager's treatment forms the first approximation of a systematic theory [5] which, however, converges only if

$$\mu R_0 < kT \quad . \quad . \quad . \quad . \quad . \quad (13)$$

i.e., if $\mu R_0/kT < 1$. This quantity can, therefore, be neglected in a first approximation which thus leads to Onsager's theory. The next approximation leads to results in between Debye's and Onsager's theories, but the exact value depends on the structure of the material; (13) is also the condition which restricts the applicability of the methods described in this paragraph. These methods imply that there should be no saturation in the polarisation near a dipole. It can be shown [5] that this again leads to (13).

Inserting R_0 from (8) and replacing $4\pi\mu^2 n/3kT$ by $(\epsilon - 1)(2\epsilon + 1)/3\epsilon$ according to Onsager's approximation (12), (13) becomes

$$\frac{2}{3}\frac{(\epsilon - 1)^2}{\epsilon} < 1, \text{ or } \epsilon < \frac{7 + \sqrt{33}}{4} \simeq 3.$$

The above condition will be replaced by

$$\epsilon < 3\epsilon_0 \quad . \quad . \quad . \quad . \quad . \quad (14)$$

if we allow for non-dipolar contributions to ϵ leading to a high-frequency dielectric constant ϵ_0. Since ϵ_0 hardly ever exceeds 3, Onsager's theory should hold only as long as $\epsilon < 9$.

4. An interesting attempt to find an expression for the dielectric constant which holds over a larger range is due to Kirkwood.[6] He showed that for liquids without the use of any approximations

$$\epsilon - 1 = \frac{\mu\bar{\mu}\, 4\pi n}{3kT}\,\frac{3\epsilon}{2\epsilon + 1} \quad . \quad . \quad . \quad . \quad (15)$$

---

[5] Fröhlich and Sack, *Proc. Roy. Soc.*, *A*, 1944, 182, 388.

[6] Kirkwood, *J. Chem. Physics*, 1939, 7, 911.

holds where $\bar{\mu}$ is the total moment of a macroscopic spherical specimen polarised by one of its dipoles which is kept in a fixed position. Equation (15) is very similar to Onsager's equation (12) but without the use of any further theory it cannot be used to calculate ϵ because $\bar{\mu}$ is an unknown function of temperature. To obtain $\bar{\mu}$ one requires a theory of the interaction between dipoles and thus one meets all the difficulties discussed before. Kirkwood's formula should, however, be very useful in the presence of strong short range forces. Equation (15) is identical with a formula which one would obtain if dipoles are assumed to associate. Suppose that on an average $\bar{\mu}/\mu$ dipoles associate forming a group of dipoles with moment $\bar{\mu}$. There are $\bar{n} = n\mu/\bar{\mu}$ such groups, and hence $\bar{\mu}^2\bar{n} = \mu\bar{\mu}n$. Thus, applying Onsager's formula to these groups leads to equation (15).

5. A treatment of a very different nature can be given for certain strongly anisotropic solids.[7] Consider, for instance, a crystal of a ketone. A ketone molecule in the crystal forms a plane zig-zag and contains one dipole whose direction is perpendicular to the axis of the molecular chain. The molecular chains are arranged in layers whose thickness is approximately equal to the chain length. In such a structure the dipoles of each layer lie in one plane, the dipolar plane, and the distance between dipolar planes usually is very large compared with the distance between neighbouring dipoles of a dipolar plane. To account for the interaction of a given dipole with all the others, it is thus possible to consider first its interaction with the other dipoles of the same dipolar plane, and then its interaction with the other dipolar planes. Now within a dipolar plane the number of dipoles in a distance between r and $r + dr$ increases only as $r dr$, and hence its contribution to the interaction is proportional to $r dr/r^3 = dr/r^2$. In contrast to the three dimensional case (where it is proportional to dr/r) this expression converges quickly, and it can be shown that it is sufficient to consider the interaction between nearest neighbours only. Problems of this type have been treated before, *e.g.* in the case of order-disorder transitions. In fact such a transition has to be expected in our present case. At low temperatures the dipoles of a dipolar plane will be arranged in an ordered pattern. With increasing temperature some of the dipoles take up different positions and at a critical temperature T_0 the dipoles are completely disordered.

The dipolar contribution to the dielectric constant of such a substance should vanish at very low temperatures. The reason for this can easily be understood. In a solid of the type considered here each dipole has two directions of equilibrium. In the ordered state each dipole is in a state of lowest energy. If a single dipole is turned into its second direction of equilibrium, its energy will be higher by an amount V. If $kT \ll V$ the probability of finding such a turned dipole will be very small $(\propto e^{-V/kT})$. Only such dipoles which can change their direction do, however, contribute to the polarisation induced by an external field.* Since their number increases with temperature the dielectric constant ϵ too will increase with temperature in contrast to the behaviour of ϵ in liquids. A detailed calculation shows that this increase continues in an accelerated way up to the transition temperature T_0 where ϵ becomes infinite. For $T > T_0$, ϵ increases approaching asymptotically the value for independent dipoles. There is no obvious relation between T_0 and the melting-point of the crystal and it may well happen that T_0 is larger than the melting-point. In this case ϵ will increase with temperature until the crystal melts. Such a behaviour has been found by A. Müller[8] on a ketone crystal. A detailed comparison between theory and experiment allows a determination of the energy V. This shows that only about half of this energy is due to dipolar interaction. The other half must be due to an interaction of a different nature.

---

[7] Fröhlich, *Proc. Roy. Soc.*, A, 1946, 185, 399, *E.R.A. Report*, L/T147, L/T156.

[8] Müller, *Proc. Roy. Soc.*, *A*, 1938, 166, 316.

* This excludes the case that the crystal has a permanent polarisation at $T = 0$.

6. No theoretical treatment has been given for other types of dipolar solids. It seems, however, that whenever a solid contains a number of distinct equilibrium positions for the dipoles it should behave qualitatively in a similar way as discussed in § 5 because our qualitative arguments hold in all these cases. Thus, either such a solid has a permanent polarisation near $T = 0$, or the dipolar contribution to the dielectric constant ϵ vanishes at $T = 0$ and increases with increasing temperature.

At sufficiently low temperatures there will be no dielectric loss of the Debye type because there is no contribution to ϵ due to the turning of dipoles. There is, however, a relatively small dipolar contribution $\Delta\epsilon$ to ϵ due to a change of the equilibrium positions of dipoles by an external field. This should be connected with a dielectric loss at very short waves of the type described by equation (4). To find this absorption one will probably approach the resonance frequency ω_0 from the side of small frequencies. In this case (4) leads to

$$\tan \phi \simeq \frac{\Delta\epsilon}{\epsilon_s} \frac{\omega\tau}{1 + \omega_0^2\tau^2}, \text{ if } \omega \ll \omega_0.$$

As a function of temperature this expression has a maximum, $\tan \phi_m = \frac{\Delta\epsilon}{\epsilon_s} \frac{\omega}{2\omega_0}$ at a temperature at which $\omega_0\tau = 1$ ($\Delta\epsilon$ and ϵ_s are independent of T). This may help to locate ω_0 if $\Delta\epsilon$ is known.

7. In conclusion it seems that the effect of dipolar interaction on the dielectric constant in the case of solids is very different from that of liquids. In the latter case it seems that it always leads to high values for ϵ connected with a negative temperature characteristic. In solids dipolar interaction primarily leads to the formation of an ordered state. This state may either be connected with a permanent polarisation or alternately the dipolar contribution to ϵ vanishes at $T = 0$ and increases with increasing temperature. Dipoles or groups of dipoles tend to orient themselves parallel to each other in the former case and anti-parallel in the latter. It thus seems to be an important task of research in dielectrics to investigate jointly the structure and the dielectric properties of solids of the type described above.

H. H. Wills Physical Laboratory,
University of Bristol,
Royal Fort, Bristol 8.

THE INFLUENCE OF HINDERED MOLECULAR ROTATION ON THE DIELECTRIC POLARISATION OF POLAR LIQUIDS.

By John G. Kirkwood.

Received 28th January, 1946.

The interpretation of dielectric polarisation from the standpoint of molecular structure depends upon the determination of the average local electric field F acting upon a representative molecule of a specimen of the dielectric substance under the influence of a macroscopic electric field E. The Lorentz estimate of the local field,

$$F = E + 4\pi P/3 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

holds where $\bar{\mu}$ is the total moment of a macroscopic spherical specimen polarised by one of its dipoles which is kept in a fixed position. Equation (15) is very similar to Onsager's equation (12) but without the use of any further theory it cannot be used to calculate ϵ because $\bar{\mu}$ is an unknown function of temperature. To obtain $\bar{\mu}$ one requires a theory of the interaction between dipoles and thus one meets all the difficulties discussed before. Kirkwood's formula should, however, be very useful in the presence of strong short range forces. Equation (15) is identical with a formula which one would obtain if dipoles are assumed to associate. Suppose that on an average $\bar{\mu}/\mu$ dipoles associate forming a group of dipoles with moment $\bar{\mu}$. There are $\bar{n} = n\mu/\bar{\mu}$ such groups, and hence $\bar{\mu}^2\bar{n} = \mu\bar{\mu}n$. Thus, applying Onsager's formula to these groups leads to equation (15).

5. A treatment of a very different nature can be given for certain strongly anisotropic solids.[7] Consider, for instance, a crystal of a ketone. A ketone molecule in the crystal forms a plane zig-zag and contains one dipole whose direction is perpendicular to the axis of the molecular chain. The molecular chains are arranged in layers whose thickness is approximately equal to the chain length. In such a structure the dipoles of each layer lie in one plane, the dipolar plane, and the distance between dipolar planes usually is very large compared with the distance between neighbouring dipoles of a dipolar plane. To account for the interaction of a given dipole with all the others, it is thus possible to consider first its interaction with the other dipoles of the same dipolar plane, and then its interaction with the other dipolar planes. Now within a dipolar plane the number of dipoles in a distance between r and $r + dr$ increases only as $r dr$, and hence its contribution to the interaction is proportional to $r dr/r^3 = dr/r^2$. In contrast to the three dimensional case (where it is proportional to dr/r) this expression converges quickly, and it can be shown that it is sufficient to consider the interaction between nearest neighbours only. Problems of this type have been treated before, *e.g.* in the case of order-disorder transitions. In fact such a transition has to be expected in our present case. At low temperatures the dipoles of a dipolar plane will be arranged in an ordered pattern. With increasing temperature some of the dipoles take up different positions and at a critical temperature T_0 the dipoles are completely disordered.

The dipolar contribution to the dielectric constant of such a substance should vanish at very low temperatures. The reason for this can easily be understood. In a solid of the type considered here each dipole has two directions of equilibrium. In the ordered state each dipole is in a state of lowest energy. If a single dipole is turned into its second direction of equilibrium, its energy will be higher by an amount V. If $kT \ll V$ the probability of finding such a turned dipole will be very small $(\propto e^{-V/kT})$. Only such dipoles which can change their direction do, however, contribute to the polarisation induced by an external field.[*] Since their number increases with temperature the dielectric constant ϵ too will increase with temperature in contrast to the behaviour of ϵ in liquids. A detailed calculation shows that this increase continues in an accelerated way up to the transition temperature T_0 where ϵ becomes infinite. For $T > T_0$, ϵ increases approaching asymptotically the value for independent dipoles. There is no obvious relation between T_0 and the melting-point of the crystal and it may well happen that T_0 is larger than the melting-point. In this case ϵ will increase with temperature until the crystal melts. Such a behaviour has been found by A. Müller[8] on a ketone crystal. A detailed comparison between theory and experiment allows a determination of the energy V. This shows that only about half of this energy is due to dipolar interaction. The other half must be due to an interaction of a different nature.

---

[7] Fröhlich, *Proc. Roy. Soc.*, A, 1946, 185, 399, *E.R.A. Report*, L/T147, L/T156.

[8] Müller, *Proc. Roy. Soc.*, *A*, 1938, 166, 316.

[*] This excludes the case that the crystal has a permanent polarisation at $T = 0$.

6. No theoretical treatment has been given for other types of dipolar solids. It seems, however, that whenever a solid contains a number of distinct equilibrium positions for the dipoles it should behave qualitatively in a similar way as discussed in § 5 because our qualitative arguments hold in all these cases. Thus, either such a solid has a permanent polarisation near $T = 0$, or the dipolar contribution to the dielectric constant ϵ vanishes at $T = 0$ and increases with increasing temperature.

At sufficiently low temperatures there will be no dielectric loss of the Debye type because there is no contribution to ϵ due to the turning of dipoles. There is, however, a relatively small dipolar contribution $\Delta\epsilon$ to ϵ due to a change of the equilibrium positions of dipoles by an external field. This should be connected with a dielectric loss at very short waves of the type described by equation (4). To find this absorption one will probably approach the resonance frequency ω_0 from the side of small frequencies. In this case (4) leads to

$$\tan \phi \simeq \frac{\Delta\epsilon}{\epsilon_s} \frac{\omega\tau}{1 + \omega_0{}^2\tau^2}, \text{ if } \omega \ll \omega_0.$$

As a function of temperature this expression has a maximum, $\tan \phi_m = \frac{\Delta\epsilon}{\epsilon_s} \frac{\omega}{2\omega_0}$ at a temperature at which $\omega_0\tau = 1$ ($\Delta\epsilon$ and ϵ_s are independent of T). This may help to locate ω_0 if $\Delta\epsilon$ is known.

7. In conclusion it seems that the effect of dipolar interaction on the dielectric constant in the case of solids is very different from that of liquids. In the latter case it seems that it always leads to high values for ϵ connected with a negative temperature characteristic. In solids dipolar interaction primarily leads to the formation of an ordered state. This state may either be connected with a permanent polarisation or alternately the dipolar contribution to ϵ vanishes at $T = 0$ and increases with increasing temperature. Dipoles or groups of dipoles tend to orient themselves parallel to each other in the former case and anti-parallel in the latter. It thus seems to be an important task of research in dielectrics to investigate jointly the structure and the dielectric properties of solids of the type described above.

H. H. Wills Physical Laboratory,
University of Bristol,
Royal Fort, Bristol 8.

THE INFLUENCE OF HINDERED MOLECULAR ROTATION ON THE DIELECTRIC POLARISATION OF POLAR LIQUIDS.

BY JOHN G. KIRKWOOD.

Received 28th January, 1946.

The interpretation of dielectric polarisation from the standpoint of molecular structure depends upon the determination of the average local electric field F acting upon a representative molecule of a specimen of the dielectric substance under the influence of a macroscopic electric field E. The Lorentz estimate of the local field,

$$F = E + 4\pi P/3 \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

where P is the polarisation of the dielectric substance, has been widely used, and although not exact, it is a fair approximation for non-polar substances.[1] For polar liquids the Lorentz approximation fails in a spectacular manner. It leads to the prediction of electric Curie points at rather high temperatures for many polar liquids. Such Curie points have never been observed. If the Lorentz field is retained, they can be avoided only by assigning to the molecules of the liquid, dipole moments entirely inconsistent with the structure which they exhibit in the vapour phase or with any reasonable modification of this structure.

The principal cause of departure from the Lorentz field in polar liquids is the hindering effect which each molecule exerts on the rotation of neighbouring molecules relative to itself. The importance of this effect was suggested by Debye [2] and Fowler,[3] but unfortunately they attempted to combine it in an heuristic way with the Lorentz field, with which it is in fact incompatible. Hindered relative molecular rotation may be due in part to electrostatic forces, and in part to short-range intermolecular forces, for example, Van der Waals forces and the exchange repulsive forces determining molecular shape.

The first important step toward the solution of the problem of the internal field in polar liquids was made by Onsager,[4] who took account of the effect of the electrostatic part of hindered rotation within the frame of a simplified model of liquid structure in which each molecule was treated as a cavity in a statistical continuum of local dielectric constant equal to that of the fluid in bulk. While Onsager's theory no longer led to preposterous conclusions, it left much to be desired from the quantitative standpoint. More recently, Kirkwood [5] has been able to develop a formally exact theory of the dielectric polarisation of polar liquids on the basis of classical statistical mechanics. This theory takes account not only of electrostatic hindering, but also hindering due to short range intermolecular forces. When used in conjunction with suitable models of liquid structure, it provides a quantitative description of dielectric polarisation of an accuracy consistent with that of the model employed.

Before proceeding to the results of the theory and to some of its applications, it is desirable to examine the physical background of the problem of the local field. As Lorentz was aware, the contribution $E + 4\pi P/3$ is that part of the local field acting upon a molecule of fixed orientation and position, which arises from the polarisation of the dielectric exterior to a spherical surface of macroscopic radius centred on the molecule. The region interior to the spherical surface is conveniently called the Lorentz cavity, but the exterior contribution is correctly given as $E + 4\pi P/3$ only if the cavity is filled with material possessing the macroscopic polarisation P. The contribution to the local field from the material in the interior of the Lorentz cavity may in special cases, for example, a cubic crystal without lattice vibrations composed of isotropic molecules, be shown to vanish, but in general this contribution does not vanish. To neglect it entirely is equivalent to assuming that the molecule in which the field is to be calculated does not produce local departures from the macroscopic polarisation in the statistical continuum constituting its environment. It may be shown that such departures are not large in the neighbourhood of an isotropic non-polar molecule. However, in the neighbourhood of a polar molecule of fixed orientation, local departures from the macroscopic polarisation play a dominant role in determining the local field.

Onsager [4] perceived that if the Lorentz cavity is allowed to shrink to the surface of a polar molecule of fixed orientation, it must be treated as a real cavity of low dielectric constant in a statistical continuum of high dielectric constant rather than as a region filled with material possessing the macrosopic polarisation. By treating the molecule as a cavity of dielectric constant equal to the square of the optical refractive index of the liquid and its environment as a structureless continuum of uniform dielectric constant equal to that of the liquid in bulk, Onsager was able to formulate a physically sensible theory of dielectric polarisation on the basis of simple electrostatic theory.

---

[1] Kirkwood, *J. Chem. Physics*, 1936, **4**, 592; Van Vleck, *J. Chem. Physics*, 1937, **5**, 556.

[2] Deybe, *Physik. Z.*, 1935, **36**, 100.

[3] Fowler, *Proc. Roy. Soc.*, *A*, 1935, **149**.

[4] Onsager, *J. Am. Chem. Soc.*, 1936, **58**, 1486.

[5] Kirkwood, *J. Chem. Physics*, 1939, **7**, 911.

The theory of Kirkwood [6] is quite general and is not based upon a specific model of liquid structure as is the theory of Onsager. By means of statistical mechanics the dielectric constant of a polar liquid is related to the molecular dipole moment and the total moment, induced in the absence of an external field as the result of hindered relative molecular rotation, by a representative molecule of fixed orientation in a spherical region surrounding it, of radius large relative to the range of intermolecular forces. The latter moment may be regarded, if so desired, as the total moment of the content of a Lorentz cavity centred on a molecule of fixed orientation in the absence of an external field. In quantitative form the theory leads to the following expression for the dielectric constant ϵ of a polar liquid,

$$(2\epsilon + 1)(\epsilon - 1)/9\epsilon = P_0/v$$

$$P_0 = \frac{4\pi N}{3}[\alpha + \mu \cdot \bar{\mu}/3kT] \quad . \quad . \quad . \quad . \quad (2)$$

where N is Avogadro's number, α the optical polarisability, μ the molecular dipole moment and $\bar{\mu}$ the sum of the molecular dipole moment and the moment induced as the result of hindered rotation by a representative molecule of fixed orientation in the spherical region surrounding it, of radius large relative to the range of intermolecular forces. As has been remarked, hindered molecular rotation may be due in part to electrostatic forces and in part to short-range intermolecular forces. Formally, the magnitude of $\mu \cdot \bar{\mu}$ depends upon the distribution function for molecular pairs, $\rho^{(2)}$, which specifies the average density of molecules at a specified distance and in a specified orientation relative to an arbitrary molecule of the fluid. When averaged over relative orientations, $\rho^{(2)}$ becomes identical with the radial distribution function which may be determined from the angular distribution of X-rays scattered by the liquid. With the use of the pair distribution function, $\mu \cdot \bar{\mu}$ may be expressed in the form,

$$\mu \cdot \bar{\mu} = g\mu^2$$

$$g = 1 + \iint \cos \gamma \, \rho^{(2)} \, dv \, d\omega \quad . \quad . \quad . \quad . \quad (3)$$

where μ is the magnitude of the molecular dipole moment, γ is the angle between the dipole moments of a representative pair of molecules, and the integration is to be extended over all relative orientations and configurations of their centres of gravity, within the sphere exterior to which the local dielectric constant may be assumed to have effectively attained the macroscopic value.

The determination of the pair distribution function $\rho^{(2)}$ requires a detailed theory of the liquid state. Although such a theory is in principle provided by statistical mechanics, the details have not yet been developed for polar liquids. Until this is accomplished, the applications of the theory of dielectric polarisation have to be based on approximate models of liquid structure. The Onsager model, in which the statistical environment of a molecule is regarded as a continuum of uniform dielectric constant, leads to a value of g equal to unity. Examination of experimental data shows, however, that g, conveniently designated as the correlation parameter, deviates significantly from the value, unity, and the deviation may be either positive or negative. Positive deviations result when short-range hindering torques favour parallel orientation of the dipole moments of neighbouring molecules. Negative deviations result when the hindering torques favour antiparallel orientation.

It has been found that reasonably reliable applications of the theory may be based upon rough quasi-crystalline models of local liquid structure, in which each molecule possesses a co-ordination shell of z nearest neighbours beyond which orientational correlation does not extend. The co-ordination number z may be determined from the area under the first peak of the radial distribution function in liquids for which X-ray scattering data are available. The model must be supplemented by assumptions concerning the torques hindering the relative rotation of neighbour molecules, suggested by their structure. For the model described, the correlation parameter g takes the form,

$$g = 1 + z\,\overline{\cos\gamma}$$

$$\overline{\cos\gamma} = \int \cos\gamma \exp(-W/kT)\,d\omega \quad . \quad . \quad . \quad (4)$$

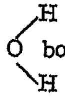
where the integral defining $\overline{\cos\gamma}$ extends over all relative orientations of a neighbour pair, and W is the potential of average torque hindering their relative rotation, with a constant term so adjusted to normalise the integral of $\exp[-W/kT]$ to unity. If W is expanded in a series of spherical harmonics and complex exponentials of the five angles specifying the relative orientation of a molecular pair, $\overline{\cos\gamma}$ may be expressed as a rather complicated function of the potential barrier heights associated with the several degrees of rotational freedom. The result is not very useful unless W possesses simplifying elements of symmetry. If, for example, W is assumed to depend only on the angle between the dipole moments of the pair and a two-term Fourier series is used, $\overline{\cos\gamma}$ takes the simple form,

$$\overline{\cos\gamma} = L(W_0/2kT) \quad . \quad . \quad . \quad . \quad (5)$$

where L is the Langevin function and W_0 is the height of the average potential barrier.

When relative molecular rotation is strongly hindered, as in the case of hydrogen-bonded liquids such as water and the alcohols, further simplifications in the model of local liquid structure, which are physically plausible, may be made. The dielectric constants of water and the alcohols have been estimated by Oster and Kirkwood [6] on the basis of models of neighbour structure admitting free rotation around hydrogen bonds, but no bending of these bonds. In the case of water, a modified Bernal-Fowler model of this type leads to the expression,

$$g = 1 + z\cos^2(\beta/2) \quad . \quad . \quad . \quad . \quad (6)$$

where β is the  $\mathrm{O}\!\begin{smallmatrix}\diagup\mathrm{H}\\ \diagdown\mathrm{H}\end{smallmatrix}$ bond angle equal to 105°, and z is the number of neighbours in the first co-ordination shell, differing slightly from four due to superposition on the tetrahedral structure of contributions from more closely packed structures. The co-ordination number z is determined from the area under the first peak of the experimental radial distribution function. Properly, $\overline{\cos\gamma}$ should not be taken as equal to $\cos^2(\beta/2)$ except for the tetrahedral structure, but for small departures from the co-ordination number, four, this assumption cannot lead to serious error, and is better than neglecting such departures altogether. The dipole moment of a water molecule in the liquid phase differs from its value in the vapour phase due to polarisation by the electrostatic field produced by the charge distributions of its neighbours. This effect is estimated to be approximately fifteen per cent. The theoretical values of the dielectric constant of liquid water based upon Eqs. (2) and (6) are compared with the experimental values at several temperatures in Table I. The discrepancies between the theoretical and experimental values are remarkably small in view of the rather oversimplified model of neighbour structure employed in the calculations.

---

[6] Oster and Kirkwood, *J. Chem. Physics*, 1943, 11, 175.

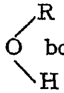
TABLE I.—DIELECTRIC CONSTANT OF WATER.

t (°C.).	ϵ (calc.).	ϵ (obs.).
0	84·2	88·0
25	78·2	78·5
62	72·5	66·1
83	67·5	59·9

Estimates of the dielectric constants of the aliphatic alcohols have been made on the basis of a similar model. It has been concluded from X-ray scattering data that chainwise co-ordination through hydrogen bonds exists in the liquid alcohols. On the assumption that orientational correlation is appreciable only between a molecule and the members of the chain of which it is a part, the correlation parameter g becomes,

$$g = 1 + 2f\cot^2(\beta/2)$$

$$f = \frac{(\mu_H + \mu_R)(\mu_H + \mu_R\cos\beta)}{[\mu_H^2 + \mu_R^2 + 2\mu_H\mu_R\cos\beta]} \quad . \quad . \quad . \quad . \quad (7)$$

where β is the  $\overset{R}{\diagup}O\overset{}{\underset{H}{\diagdown}}$ bond angle in the alcohol ROH and μ_H and μ_R are the moments of the respective bonds, O—H and O—R. The theoretical dielectric constants of series of aliphatic alcohols based upon this model are compared with the experimental values in Table II. The agreement with experiment is evidently satisfactory in view of the roughness of the neighbour structure model which was employed in the estimates.

Eqs. (2) and (3) may be employed in the manner described to predict the dielectric constants of polar liquids on the basis of simplified models of liquid structure. The precision of such calculations is limited by the degree of approximation to which the model represents the molecular pair distribution function $\rho^{(2)}$. The theory may, on the other hand, be employed in another way. From observed dielectric constants and values of vapour phase dipole moments conventionally corrected to the liquid phase, Eqs. (2) and (3) yield values of the correlation parameters g for polar liquids. The departure of g from the value unity is a measure of the degree of hindered relative molecular rotation arising from short-range intermolecular forces. From general statistical mechanical considerations it is evident that the value of g must be qualitatively related to the rotational entropy of the liquid. Examination of the values of g calculated from observed dielectric constants brings out this relationship rather clearly. Those polar substances possessing values of g departing but slightly from unity exhibit the thermo-dynamic properties of "normal" liquids. Those substances possessing values of g departing significantly from unity exhibit the properties of "abnormal" or "associated" liquids.

TABLE II.—DIELECTRIC CONSTANTS OF THE ALIPHATIC ALCOHOLS.

	(Calc.).	(Obs.).
Methyl alcohol . .	29·2	32·8
Ethyl alcohol . .	21·3	24·6
n-Propyl alcohol .	17·3	19·5
n-Butyl alcohol . .	14·6	18·0
n-Amyl alcohol . .	12·8	15·8

The theory in its present form is of course an equilibrium theory and does not describe anomalous dispersion and dielectric loss in alternating electric fields. It has not yet been possible to formulate the generalisation of the theory necessary for the description of dielectric loss, but it is hoped that this may be accomplished by an analysis of the relative rotatory Brownian motion of molecular pairs.

University of Chicago,
Chicago, 37,
Illinois, U.S.A.

THEORY OF THE CRYSTALLINE FIELD IN SOLID AND LIQUID POLAR DIELECTRICS.

BY EDMOND BAUER AND DANIEL MASSIGNON

One of us established in 1943 a theory which may have enabled the phenomena occurring in solid and liquid polar dielectrics to be better understood than it was in older theories.[1] This theory is a generalisation of ideas published in 1935 by Debye and Fowler.[2]

We assume that the action on a dipole of a lattice and, for a liquid, of neighbouring molecules, can be represented by a " crystalline field ". In opposition to Debye's and Fowler's assumptions, this field is not generally uniform for the following reasons.

(1) The interactions between molecules are not entirely of electrostatic origin (Van der Waal's forces, steric hindrances, etc.).

(2) In calculating the electrostatic forces themselves, we must take into consideration the charge distribution in the molecules : since the intermolecular distances are of the same order as the distance separating the poles of a dipole.

The energy of orientation of a molecule in the crystalline field must be obtained from the most general form of a function $W(\theta, \phi)$, of the angles which determine that orientation.

This potential energy may be such that the molecular dipole would possess several positions of stable equilibrium, separated by potential barriers. It is possible to perform the calculations for the general case, and to find formulæ giving the relaxation time and the dielectric constant. Before doing this, we must make an assumption similar to that made in the theory of reaction rates : we suppose that statistical equilibrium is attained very rapidly in every region bounded by a potential barrier ; and that the relaxation effects arise from the passage of dipoles across this barrier. It is therefore necessary, when we consider dielectric polarisation to make the distinction not only between the electronic and the atomic polarisation, and the orientation polarisation, but also in the latter between an *instantaneous* orientation polarisation (in every potential hollow) and a *retarded* polarisation (involving transitions between hollows).

Towards higher frequencies, experiments on electrical dispersion always show a tailing-off in the curve, which is very much at variance with the theory of Debye ; but the experiments are in agreement with the above method of treatment.

The word " instantaneous " must not be used in an absolute sense ; compared to the period of the electric field (even for centimetre waves), the time which is necessary for the instantaneous polarisation to reach its maximum value is very short. This polarisation is the result of the perturbation of the crystalline field by the external field (perturbation of the first order) ; its magnitude is related to the constants of the quasi-elastic forces that give the first approximation to the crystalline field. Thus it is intimately connected with the libration frequency of the dipolar molecules in the crystal (or in the liquid). We have here a phenomenon of dispersion similar to the usual optical dispersion.[3]

---

[1] Bauer, *Cahiers de Physique*, 1944, **20**, 1 ; 1944, **21**, 37 ; 1945, **27**, 33.

[2] Debye, *Physik. Z.*, 1935, **36**, 100 ; Fowler, *Proc. Roy. Soc., A*, 1935, **149**, 1.

In that theory it is difficult to calculate the real action of the external field on the dipoles within the dielectric.

Our earlier work was based on the Mosotti-Lorentz-Debye theory (internal field equal to $E + \frac{4}{3}\pi I$), which is a rough approximation. A perhaps slightly better approximation, with errors in the opposite direction, may be obtained by a modified Onsager theory.

In the following two sections we shall sum up the results of both calculations, which are not essentially different.

2. If one combines the Lorentz-Debye hypotheses with that of the crystalline field, *i.e.*, that of " hindered rotation ", one can deduce a general formula for the molar polarisation of orientation P_μ. In the case of a powder, the result is :

$$P_\mu = (P_\mu)_{\text{gas}} \cdot \left(1 - \frac{\Phi^2 + A^2 + B^2}{\sigma^2}\right) \quad . \quad . \quad . \quad (1)$$

in which we assume that the electric moment of the molecule is the same in a condensed phase as it is in a gas. In formula (1), we have

$$\sigma = \int e^{-\frac{W}{kT}} \cdot d\Omega \text{ (configuration partition function)} \quad . \quad (2)$$

$$\Phi = \int \cos\theta e^{-\frac{W}{kT}} d\Omega, \quad A = \int \sin\theta \cos\Phi e^{-\frac{W}{kT}} d\Omega, \quad B = \int \sin\theta \sin\Phi e^{-\frac{W}{kT}} d\Omega.$$

These equations hold for any form of the function W ; but it is possible to derive approximate formulæ by making simple assumptions concerning the shape of the crystalline field. Thus the Debye-Fowler's formula is a particular case of (1) ($W = -\beta\cos\theta$) ; one can develop the calculations in a more general case, in which the crystalline field is axially symmetric, with equilibrium positions having different energy minima at $\theta = 0$ (corresponding, say, to $W = 0$), and $\theta = \pi$ (having $W = U > 0$) and a barrier of potential W_m at some intermediate angle θ_m.

In that case, provided that the temperature is sufficiently low, and W_m and U are high enough, the molar polarisation of orientation will be independent of temperature :

$$P_\mu = \frac{8\pi}{9} \cdot \frac{N\mu^2}{\beta} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Here β is the elastic constant that determines the libration frequency. This formula is analogous to that obtained in the theory applying to cases of paramagnetism independent of temperature. It gives in all cases, in first approximation, the value of the *instantaneous* orientation polarisation.

This model explains also the relaxation phenomena. The instantaneous polarisation is caused by the changes (due to the external field) of the statistical molecular distribution in both regions D_1 and D_2 separated by the potential barrier. The retarded polarisation is the result of the passage of the dipole axes over the potential barrier. In a general model, in which the top of the barrier is an arbitrary curve (Λ), on the sphere of radius 1, the formula obtained for the relaxation time is :

$$\frac{1}{\tau} = \sqrt{\frac{kT}{2\pi I}} \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2}\right) \cdot e^{-\frac{W_m}{kT}} L \quad . \quad . \quad . \quad . \quad (4)$$

---

[3] In *Nature*, 1946, **157**, 478, H. Fröhlich recently introduced a damping term into the instantaneous polarisation ; he derived from it the optical width of the libration band, and interpreted thus the residual effects of electrical absorption which go with the tailing-off of the dispersion curve.

Cf. also eq. 4 of his paper at the present discussion.

where I is the moment of inertia of the dipole molecule,

σ_1, σ_2 are the configuration partition functions corresponding to each region D_1 and D_2,

L is the " effective length " of the potential barrier :

$$Le^{-\frac{W_m}{kT}} = \int_{\Lambda} e^{-\frac{W}{kT}}(d\theta + \sin\theta \, d\Phi) \quad . \quad . \quad . \quad (5)$$

The logarithm of τ is, to a good approximation,[4] a linear function of $1/T$. Moreover, the preceding theory is a particularly simple case of the general theory of reaction rates developed by Marcellin,[5] Eyring, and others.

It would be too long to write here the general equation giving the value of the retarded polarisation. But, in the simple case of axial symmetry, all the measurable quantities depend only on the height W_m, and on the angle θ_m of the potential barrier, and on U, the secondary minimum of potential energy. When we apply this model to ice, putting $\theta_m = 120°$, which corresponds roughly to the Bernal structure, we find :

$$(P_\mu)_{\text{inst.}} = (P_\mu)_{\text{gas}} \cdot \frac{3kT}{gW_m} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

$$(P_\mu)_{\text{ret.}} = (P_\mu)_{\text{gas}} \cdot \frac{4}{3} e^{-\frac{u}{kT}} \frac{W_m}{W_m - U}, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

$$\tau = (1/1{\cdot}74\, g.) \cdot \frac{e^{(W_m - U)/kT}}{(W_m - U) + (W_{m/3})e^{\frac{u}{3kT}}} . \quad . \quad . \quad (8)$$

the calculations being made by replacing the two slopes of the potential mountain by the simple algebraic curves $W_1 = W_1(x)$ and $W_2 = W_2(x)$, where $x = \cos\theta$; and,

$$g = \begin{cases} 1 \text{ linear approximation} \\ 2 \text{ parabolic approximation} \\ 3 \text{ cubic approximation} \end{cases}$$

the curves W_1 and W_2 having, in the two last approximations, the same horizontal tangent for $\theta = \theta_m$.

The measurement of the dielectric constant on one hand, and of the relaxation time on the other, give us two different means of calculating the constants W_m and U. The results of these two methods are in fairly good agreement, and of a reasonable order of magnitude.

The instantaneous polarisation of orientation makes it possible to calculate the libration frequency of H_2O molecules in ice ; instead of 600 cm.$^{-1}$, we find 360 cm.$^{-1}$, which is too little, but of the right order of magnitude.

Changing from ice to water, the dielectric constant and the frequency of libration (500 cm.$^{-1}$) change but little. From this, we conclude that the form of the crystalline field is not sensibly altered. This is in agreement with short range order of the quasi-crystalline structure of water. In contrast, the relaxation time passes sharply from 10^{-6} to 10^{-11} sec. This fact is immediately explained if it is assumed that the relaxation time in the liquid is a measure of the life of the quasi-crystalline cells : before the dipoles have time to cross the potential barrier the latter breaks down and another barrier is formed. We do not know the exact size of those cells ; they may be formed by one central molecule and its four nearest neighbours.

The same theory allows a discussion of the dielectric properties of crystals near a transition point of the second order (HBr) : the considerable increase of relaxation time τ, when the temperature draws near the transition temperature, and its sudden decrease at this very point show us the gradual intervention of a potential barrier and its sudden breakdown.

A similar discussion of dielectric phenomena in alums and ammonium iodide is possible. In alums, the water molecules librate in the field of ions, and not of dipoles: this simple fact explains the difference between substances like HBr and the alums, where are no second order transition points.

---

[4] In a previous, less detailed investigation, Eyring and his collaborators found a similar result (Eyring, *The Theory of Rate Processes*, 1941, p. 544).

[5] R. Marcelin, *Ann. de Physique*, 1915, 3, 120 ; Eyring, *Chem. Rev.*, 1935, 17, 65.

3. In his theory, Onsager[6] divides the internal field in two, a *cavity field* which alone orients the molecules, and a *reaction field* which only increases their electric moment and always remains parallel to the dipole.

It seems impossible to admit this assumption. In fact, it is necessary to distinguish two components in the reaction field, one due to electronic and atomic polarisations,[7] which follows the dipole in its oscillations, and another due to orientation polarisation. The latter is subject to relaxation phenomena characterised by a period of time τ very much longer than the period of oscillation (10^{-7} sec. instead of 10^{-13} sec. for ice).

A modified Onsager theory may be combined to the theory of the crystalline field. The calculations are simple only when the crystalline field is axially symmetric, with two equilibrium positions separated by a potential barrier, a model already used in the preceding section. In that case, the second component of reaction field must remain in each region parallel to the mean orientation of the dipole in this region.

Thus the reaction field contributes to the quasi-elastic force corresponding to the libration frequency, and may be considered as a component of the crystalline field; but the orientation of this component would change in the external field.

So we shall suppose that in the D_1 and D_2 regions, we have for both the reaction fields:

$$\vec{R}_1 = r\overrightarrow{\overline{M}_1 j_1} \quad \text{and} \quad \vec{R}_2 = r\overrightarrow{\overline{M}_2 j_2}, \qquad . \quad . \quad . \quad (9)$$

where r is a constant which can be made equal to that of Onsager

$$= \frac{2(\epsilon - 1)}{2\epsilon + 1} \cdot \frac{4\pi}{3} \cdot \frac{N'}{V},$$

$\overline{M}_1$ and $\overline{M}_2$ are the mean components of the electric moment projected on to the mean directions $\vec{j}_1$ and $\vec{j}_2$ of the dipole in each region.

Nothing is altered in the qualitative results. The formula (4), giving the relaxation time, remains valid. Whereas in the above paragraph we found a value of U probably too low, here we find one which is probably too high; but the calculation of the libration frequency gives a better result: when we take into account the fact that the electric moment of the H_2O-molecule of ice is about 20 % higher than in the gas, because of molecular interaction, we find by our new theory that the libration frequency of ice is between 450 and 500 cm.$^{-1}$.

With the two methods of calculation (sections 2 and 3) we find a rather large average number of molecules H_2O which are in the potential hollow about $\theta = \pi$, *i.e.* 3 % in the first theory and a little less in the second.[8]

There remains one difficulty; the calculation of the reaction field near the potential barrier gives two different values, according as the dipole crosses it from the region D_1, or from the region D_2. On the other hand, the dipole remains a relatively short time in the D_2 region. Our hypotheses on the reaction field are right probably only in the region D_1 near the position $\theta = 0$, where, in fact, we find the largest number of dipoles. Possibly it will be necessary in the other cases to introduce corrections similar to those which Fröhlich[9] has put forward for a simpler model.

Laboratoire de Chimie Physique,
11 *rue Pierre Curie,*
Paris (*Ve*),
France.

---

[6] Onsager, *J.A.C.S.*, 1936, 58, 1486.

[7] Bauer and Magat, *J. Physique*, 1938, 9, 319.

[8] This fact may perhaps explain the abnormal diffusion spots of X-rays observed by Mrs. Lonsdale in ice.

[9] Fröhlich, *J. Inst. Elect. Engin.*, 1944, 91, 486.

A CONTRIBUTION TO THE THEORY OF THE INTERNAL ELECTRICAL FIELD.

By C. J. F. Böttcher.

Received 3rd March, 1946.

When a homogeneous electrical field $\vec{E}$ prevails in a dielectric, the internal electrical field at the place of a particle of the dielectric can be defined as $\vec{F} = \vec{m}/\alpha$ in which α is the mean polarisability of the particle and $\vec{m}$ is the mean electric moment in the direction of $\vec{E}$, due to its polarisability only.

It follows from this definition that when we suppose the particle to be removed after fixing the distribution of the charges in its surroundings the mean value of the electrical field that would be measured in the cavity obtained is equal to the internal field $\vec{F}$.

In the classical method of calculating $\vec{F}$, as it was finally arrived at by Lorentz,[1] a sphere is constructed around the point A where the field $\vec{F}$ must be calculated, its radius being large in proportion to the distances between the molecules. The result of the calculation is:

$$\vec{F} = \vec{E} + \frac{4}{3}\pi\vec{P} + \vec{F}_1 = \frac{\epsilon + 2}{3}\vec{E} + \vec{F}_1 \quad . \quad . \quad . \quad (1)$$

in which $\vec{P}$ is the total electrical moment per cc. (equal to $E\,(\epsilon - 1)/4\pi$), ϵ the dielectric constant and $\vec{F}_1$ the contribution to $\vec{F}$ resulting from the particles inside the large sphere.

Lorentz calculated $\vec{F}$ only in the special case of a cubic crystal; for liquids and gases we shall carry out the calculation of $\vec{F}_1$ in a somewhat different way from Lorentz, dividing the large sphere into a very great number of small volume elements of equal size. The mean value in time of the components $(p_x)_i$, $(p_y)_i$ and $(p_z)_i$ of the dipole moment $\vec{P}_i$ that belongs to a definite element i with co-ordinates x_i, y_i, z_i is called $\overline{(p_x)_i}$, $\overline{(p_y)_i}$ and $\overline{(p_z)_i}$. Now $\vec{F}_1$ will be given by

$$F_1 = \sum_i \overline{(p_x)_i}\left(\frac{3x_i^2 - r_i^2}{r_i^5}\right) + \sum_i \overline{(p_y)_i} \cdot \frac{3x_i y_i}{r_i^5} + \sum_i \overline{(p_z)_i} \cdot \frac{3x_i z_i}{r_i^5} \quad . \quad (2)$$

when the x-axis has the direction of $\vec{E}$ and A is the centre of the co-ordinate system.

1. If A is a fixed point in a gaseous or liquid dielectric this point will, in the course of time, be inside a molecule at certain times and outside the molecule at other times. Consequently $(p_x)_i$ will have the same value for all the volume elements and so will $(p_y)_i$ and $(p_z)_i$. Thus:

$$F_1 = \overline{p_x} \cdot \sum_i \frac{3x_i^2 - r_i^2}{r_i^5} + \overline{p_y} \cdot \sum_i \frac{3x_i y_i}{r_i^5} + \overline{p_z} \cdot \sum_i \frac{3x_i z_i}{r_i^5} \quad . \quad (3)$$

---

[1] Lorentz, *Theory of Electrons*, Leipzig, 1909, pp. 138, 306.

Lorentz showed already that by summarising over all points of a sphere or a great number of points that are divided symmetrically over a sphere:

$$\sum_1 \frac{3x_1^2 - r_1^2}{r_1^5} = 0 \qquad \sum_1 \frac{3x_1y_1}{r_1^5} = 0 \qquad \sum_1 \frac{3x_1z_1}{r_1^5} = 0 \qquad . \quad (4)$$

So (3) turns into $F_1 = 0$.

Instead of taking for A a definite point of the dielectric and averaging F over the time we should have obtained the same result by taking the mean value of F over all points of the dielectric at a certain moment. Thus for an arbitrary dielectric the mean value in space $\bar{\bar{F}}$ of F is equal to:

$$\bar{\bar{F}} = \frac{\epsilon + 2}{3} E \quad . \quad . \quad . \quad . \quad . \quad (5)$$

So even in the case of an arbitrary dielectric the classical formula for the internal field is of use in some respect, in that it represents the average in space of the internal field.

2. If, on the other hand, A is a fixed point of a particle, *i.e.* a point that moves with the molecule during the time, there will be a correlation between the position of a certain volume element in relation to A and the time average of the dipole moment belonging to this volume element, for the inhomogeneous field caused by the particle to which A belongs, contributes in this case to the state of polarisation of the volume element under consideration. This fact is also the starting point of Kirkwood's method[2] for a statistical treatment of the problem. The method introduced by Lorentz is here not suitable for the computation of $\vec{F}$ at the place of a molecule in a gas or a liquid. Another method will be given to calculate $\vec{F}$ in this case and it will be shown that in a mixture F has a different mean value at the place of the different kinds of particles.

We consider a dielectric, for the time being, with only one kind of particles. Formerly[3] we inferred already the following equation for the internal field:

$$F = E_r + R\,\overline{\cos\phi} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

in which $\overline{\cos\phi}$ is the mean value of the cosine of the angle between the dipole axis and the direction of $\vec{E}$. We call $\vec{R}$ the reaction field of the permanent dipole and $\vec{E}_r$ the directing field. Formula (6) can be interpreted as follows:

a rigid permanent dipole polarises its surroundings which on their part cause a reaction field $\vec{R}$ at the place of the dipole, proportional to the dipole moment $\vec{\mu}$ and thus given by

$$\vec{R} = f\vec{\mu} \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

f being a constant.

For a polarisable dipole with polarisability α_0 in the direction of the dipole axis, (7) turns into:[4]

$$\vec{R} = f(\vec{\mu} + \alpha_0\vec{R}) \quad \text{or} \quad \vec{R} = \frac{f}{1 - f\alpha_0}\vec{\mu} \quad . \quad . \quad . \quad (8)$$

When a homogeneous field E prevails in the dielectric the dipole is directed and thus the reaction field will contribute to the internal field with an amount of $R\,\overline{\cos\phi}$. Onsager [5] showed already that the part of the internal field, due to the reaction field of the permanent dipole has no directing influence on the dipole and so we will call the remaining part of $\vec{F}$ the directing field $\vec{E}_r$.

---

[2] Kirkwood, *J. Chem. Physics*, 1936, 4, 592.

[3] Böttcher, *Physica*, 1938, 5, 635.

[4] Böttcher, *ibid.*, 1938, 5, 635; *Rec. trav. chim.*, 1943, 62, 119.

As with the usual values of E no saturation occurs, Langevin's well-known formula can be used :

$$\overline{\cos\phi} = \frac{\mu}{3kT}E_r \quad \ldots\ldots \quad (9)$$

in which T is the absolute temperature.

Combination of (6) and (9) leads to

$$\vec{F} = \frac{3kT + \mu R}{3kT}\vec{E}_r \quad \ldots\ldots \quad (10)$$

Incidentally we note [4] that (10) leads immediately to the semi-empirical relation of van Arkel and Snoek [6] for the dielectric constant of dipole liquids

Now $\vec{E}_r$ can also be defined in the following way: Remove the permanent dipole of the molecule without changing its polarisability and let the surrounding dielectric adapt itself to the new situation. The internal field at the place of this "maimed" molecule is then equal to $\vec{E}_r$. With the aid of this definition it is possible to calculate $\vec{E}_r$ in a simple way.

If in the above prescription we should not only have removed the permanent dipole but also the polarisability of the molecule we should have formed a spherical cavity in the dielectric, but now not—as in the definition of the internal field—fixing the distribution of charges in its surroundings but by letting themselves adapt to the new situation. This is a well-known electrostatic problem; its solution is that in the cavity arises a homogeneous field $\vec{E}_h$ equal to

$$\vec{E}_h = \frac{3\epsilon}{2\epsilon + 1}\vec{E} \quad \ldots\ldots \quad (11)$$

Now we have defined the directing field in such a way that the polarisability of the molecule must still be considered. The field $\vec{E}_r$ causes a dipole $\alpha\vec{E}_r$ and its reaction field is $f\alpha\vec{E}_r$. So $\vec{E}_r$ will be given by the implicite equation :

$$\vec{E}_r = \vec{E}_h + f\alpha\vec{E}_r \quad \text{or} \quad \vec{E}_r = \frac{\vec{E}_h}{1 - f\alpha} = \frac{1}{1 - f\alpha}\cdot\frac{3\epsilon}{2\epsilon + 1}\vec{E} \quad . \quad (12)$$

Substitution of the equations (8) and (12) for $\vec{R}$ and $\vec{E}_r$ in (10) leads to

$$\vec{F} = \frac{3kT + \dfrac{f}{1 - f\alpha_0}\mu^2}{3kT}\cdot\frac{1}{1 - f\alpha}\cdot\frac{3\epsilon}{2\epsilon + 1}\cdot\vec{E} \quad \ldots \quad (13)$$

The factor f of the reaction field was calculated by Onsager [5] who found

$$f = \frac{1}{r^3}\cdot\frac{2\epsilon - 2}{2\epsilon + 1} \quad \ldots\ldots \quad (14)$$

in which r is the radius of the particle and ϵ the dielectric constant of its surroundings.

In a mixture the derivation of formula (13) remains true, but in this case α, r, f and μ refer to the considered kind of particle. So it results that in a mixture, F has a different value at the place of the different kinds of molecules. This is also expressed by equation (6): in a diluted solution of a dipole compound in a dipole-free solvent the field at the place of a dipole molecule is much greater than at the place of a molecule of the solvent as a consequence of the term $R\,\overline{\cos\phi}$. For instance in a solution of nitrobenzene in benzene (1 g. mol. : 10 g. mol.) at $T = 300°$ K. we find:

---

[5] Onsager, *J. Am. Chem. Soc.*, 1936, 58, 1486.

[6] van Arkel and Snoek, *Physik. Z.*, 1932, 33, 662; 1934, 35, 187.

$$F_{\text{benzene}} = 1{\cdot}6\,E \text{ and } F_{\text{nitrobenzene}} = 11{\cdot}2\,E.$$

We have shown already[7] that formula (13) for the internal field complies with the demand that the mean value in space shall be equal to $E\,.\,(\epsilon + 2)/3$.

From the foregoing theory on the internal field the deviations from the formulæ of Lorentz-Lorenz, Clausius-Mosotti and Debye can be accounted for, *e.g.* the dependence on the pressure of the molecular polarisation of gases can be explained[8] as well as the dependence on the concentration of the molecular refraction of electrolytes in aqueous solutions.[9] From the latter dependence the radii of the ions can be calculated. Formula (5) can be used for the calculation of the dielectric constant of a crystal from that of a crystalline powder.[10]

The University,
Leyden.

[7] Böttcher, *Physica*, 1942, **9**, 937.

[8] Böttcher, *ibid.*, 1942, **9**, 945.

[9] Böttcher, *Rec. trav. chim.*, 1943, **62**, 325, 503; 1946, **65**, 14, 19, 39, 50

[10] Böttcher, *ibid.*, 1945, **64**, 47.

A COMPARISON OF DEBYE'S AND ONSAGER'S THEORIES OF THE DIELECTRIC CONSTANTS OF CONCENTRATED POLAR MEDIA.

By F. C. Frank.

Received 4th March, 1946.

The problem of the way in which the dielectric constant of a condensed medium is related to the properties of its individual molecules is a complicated one, made more difficult by subtle differences in various authors' approaches to the same problem. The oldest problem, concerned with the dielectric constant or refractive index of a medium composed of non-polar molecules, can now be regarded as settled: we still await a completely satisfactory treatment of the case where the molecules possess permanent dipoles.

Practically all treatments adopt the device of singling out one central molecule, finding the appropriately averaged influence on it of all the rest, and finally making the solution self-consistent in recognising that these surrounding molecules must have the same average properties and behaviour as the central one. An elaboration of this technique consists in describing an imaginary surface at an arbitrary distance from the central molecule and discussing separately the influence on it of molecules outside and inside this surface.

When the molecules are non-polar we have only to consider fields due to the measuring field E, and the problem is one of deciding how the average field on our central molecule depends on this. If the molecule were situated in a pipe drilled parallel to the field, it would be $E_p = E$; in a crack or saw-cut perpendicular to the field it would be $E_c = \epsilon E$ where

ϵ is the dielectric constant of the medium as a whole; while in a spherical cavity it would be $E_s = \frac{2}{3}E_p + \frac{1}{3}E_c = (2 + \epsilon)E/3$, if that cavity could be scooped out without disturbing the state of polarisation of the surrounding medium. This important proviso, which may be seen to be implied in, *e.g.* Debye's derivation (ref. [1], pp. 9 ff.) is omitted from the otherwise careful statement of the case in ref. [10], paragraph 1424. Mosotti concluded that the latter field, E_s, would be the field acting on his central molecule and (in 1850) derived the well-known formula

$$\frac{\epsilon - 1}{\epsilon + 2}V = \frac{4\pi N}{3}\alpha.$$

Lorenz reached the same conclusion in considering refractive indices (giving the same formula with $n^2 = \epsilon$). Clausius and Lorentz subsequently lent the weight of their authority to the conclusion, and Van Vleck has recently derived it quantum-mechanically for a restricted model. The formula is commonly designated by the names of the four first-mentioned authors, taken in pairs or together: here we shall be satisfied with that of the originator. The Mosotti formula is established, both theoretically and experimentally, at least as a very good approximation, so long as we are concerned with distortion-polarisation only.

When the medium contains polar molecules, we have other components of the electrical field to consider besides—

(*a*) that arising from the field applied for measurement.

There is (*b*) the field due to the permanent moment of our central molecule, which polarises the surrounding medium and also produces thereby a reacting field on the central molecule itself;

there is also (*c*) the resultant field arising from neighbouring polar molecules, which will not in general be zero even when we are considering a non-polar central molecule:

and there may be (*d*) non-electrostatic forces restricting molecular rotation, which can reduce the contribution of dipole orientation to the dielectric constant.

Debye's theory of permanent dipoles [1] was originally worked out for gases, in which the intermolecular interactions can generally be neglected. Its extension to dilute solutions (with extrapolation to infinite dilution to eliminate interactions between the polar molecules) was originally tentative, and tested by experiment, mainly by Williams. The Mosotti formula was used to provide a law of mixtures. Some effect of interaction between the polar molecule and its non-polar solvent medium, dependent on the dielectric constant of the latter, was expected: however, partly because of experimental inaccuracies and partly because of a restricted choice of materials, no effect was discovered, and a great many measurements of dipole moment in solution were made before Müller [5, 6] showed that these solvent effects genuinely occur. They can be adequately explained, if only approximately calculated, by the field (*b*) mentioned above, taking due account of the shape of the polar molecule. This is of great importance, determining the sign as well as the magnitude of the effect (Weigle,[7], Higasi,[8] Frank [9]).

It was recognised from the start that some modification of the theory would be required to deal with concentrated solutions or pure polar

[1] Debye, *Polar Molecules*, New York, 1929.
[2] Debye, *Physik. Z.*, 1935, **36**, 100, 193.
[3] Debye and Ramm, *Ann. Physik*, 1937, **28**, 28.
[4] Onsager, *J. Amer. Chem. Soc.*, 1936, **58**, 1486.
[5] Müller, *Physik. Z.*, 1932, **33**, 731.
[6] Müller, *ibid.*, 1933, **34**, 689.
[7] Weigle, *Helv. Phys. Acta*, 1933, **6**, 68.
[8] Higasi, *Abstr. Bull. Inst. Phys. Chem. Res., Tokyo*, 1934, 65.
[9] Frank, *Proc. Roy. Soc., A*, 1935, **152**, 171.

liquids. In 1935 Debye produced his "hindering energy" theory [2,3] in which fields (*c*) and (*d*) together are regarded as producing partial dielectric saturation, reducing the degree of orientation which can be produced by the force arising from the measuring field. This "hindering energy" appears as a new parameter, regarded as a measure of the "quasi-crystallinity" of the liquid. It is unfortunate that we have at present no accurate independent way of measuring this energy restricting rotation so as to check the theory. Debye's disregard of the field component (*b*) may be only a minor defect (though it is more serious in media of high dielectric constant than in non-polar solvents): the gravest uncertainty in the theory arises from his tacit retention of the Mosotti formula for the field (*a*). This may be right, but proof of its rightness is lacking.

Onsager [1] approached the problem afresh, avoiding the latter assumption and attempting to evaluate fields (*a*) and (*b*) directly from the laws of electrostatics. His model of the central molecule is a sphere of uniform refractive index n, with radius a, also possessing a permanent dipole moment μ_0 (the moment which would be measured in the vapour state). This sphere is surrounded immediately by a medium of uniform dielectric constant ϵ, representing the smoothed effect of the surrounding polar molecules. He derives a field acting on the central molecule which can be resolved into two components, one which we may call F_E, dependent on and parallel to the measuring field E, and the other which we may call F_μ, dependent on and parallel to the primary dipole μ_0. These have the values

$$F_E = \frac{(n^2 + 2)\epsilon}{n^2 + 2\epsilon} \cdot E.$$

$$F_\mu = \frac{2(n^2 + 2)(\epsilon - 1)}{3(n^2 + 2\epsilon)} \cdot \frac{\mu_0}{a^3}.$$

The latter component of field, exerting no couple on the molecule, is relatively unimportant. A correction to the effective dipole moment (of up to 50 %) can be derived from it, but as it only applies to the assumed model of a spherical molecule, it is of limited practical significance. The essence of Onsager's theory lies in the former component. This is only equal to the Mosotti field when $\epsilon = n^2$. Unlike the Mosotti field which in media of high dielectric constant rises proportionately to ϵ, this field soon approaches a much smaller constant value. It is on this account that the Onsager theory, without any "hindering energy" predicts smaller dielectric constants for polar liquids, with no tendency for the dielectric constant to pass to infinity at a critical temperature. It is to be noted that when $n^2 = 1$ the Onsager field component F_E becomes $E \cdot 3\epsilon/(2\epsilon + 1)$, *i.e.* the field in a genuine empty spherical cavity, unlike E_i defined above, which can be paradoxically described as the field in "a filled cavity with negligible contents". The difference arises from the quite different significance of the two imaginary cavities. Onsager's cavity contains a single molecule, whereas in expositions, *e.g.* Debye's, leading to the Mosotti result, the cavity is large and contains many molecules. The argument is that the influence of molecules outside the cavity produces Mosotti's field, while the influence of molecules inside the cavity, on the central one, can be shown to average to zero in certain simple typical cases (not involving polar molecules). Onsager holds that Debye's extension of the same theory to media containing polar molecules involves a false averaging process: that only the component F_E can exert a couple on the dipole, and that the reaction field F cannot do so because it is always parallel to the dipole. This, however, is only so if we can truly represent the surrounding polar molecules, as they respond to the field of the central dipole, as an ideal homogeneous dielectric (having the macroscopic dielectric constant of the liquid). Now the field strength at, say, 4 A. from a typical molecular dipole of 2 D. is of the order μ_0/a^3, or 10^7 volts/cm. The potential energy of a second similar dipole in this field is 6×10^{-14} ergs or $1\frac{1}{2}$ kT, at ordinary temperatures. But the dielectric "constant" of a polar liquid is only constant and isotropic for fields in which $\mu F \ll kT$. That is to say Onsager has neglected dielectric saturation which must be present and is the essence of Debye's theory. Onsager has eliminated Debye's "hindering energy," not by proof of absence or unnecessity, but by simple disregard.

Thus it is clear that neither Debye's nor Onsager's theory can hope to be a final complete theory of the dielectric constants of polar liquids (Debye's theory in any case omits our field component (*b*)). As for their merits as providing self-consistent models illustrating the essentials of reality, it is clearly a matter of taste whether one is more distrustful of the element of guesswork in Debye's theory, employing the Mosotti field outside its proven range of validity, or of a demonstrable important omission from Onsager's treatment. There must be some degree of truth in Debye's supposition of a "hindering energy" (*i.e.* that the molecular interaction field does not wholly readjust itself to accommodate any change of orientation of the particular molecule on which attention is focussed): and the truer this element of Debye's theory, the weaker the case for using Onsager's version of the inner field in place of Mosotti's. In the extreme case of a crystal of polar molecules having low dielectric constant (*e.g.* solid nitrobenzene) we cannot doubt that there are molecular interactions (in which the electrostatic fields of the dipoles are a not unimportant part), for which Debye's "Behinderungsenergie" is a sensible first approximation. Then the contribution to the whole polarisation of small changes of molecular orientation is not *very* different in mechanism from a distortion-polarisation, so that Mosotti's formula for the inner field arising from an applied measuring field is probably the best approximation we can get without doing a vastly greater amount of calculation for the specific case. (The foregoing paragraph is largely inspired by a remark made by Professor N. F. Mott.)

The Onsager theory accounts for the order of magnitude of the dielectric constant of all but a few highly associated liquids, without the introduction of any hindering energy. When this is introduced, as we saw it must be, presumably the results will be too low. The Debye theory cannot be thus confirmed against experiment, for it contains in the hindering energy what must be regarded as an arbitrary parameter so long as there is no independent way of measuring it with comparable accuracy. Debye's application of his theory to dielectric saturation in high measuring fields, with surprisingly satisfactory result in the case of water, provides a partially independent check: but water is an unsatisfactory test substance, for which Debye's theory is undoubtedly and admittedly too crude. The truth may lie in some combination of the two theories, more on the side of Onsager in some cases, and more on the side of Debye in others. It is fair to say Onsager has shown Debye's theory to be a less excellent demonstration of the quasi-crystallinity of liquids than it at first appeared. But if one believes on other grounds, as this author does, that liquids are often not very different from crystals in respect of the interactions of adjacent molecules, then one must take Debye's theory, and not Onsager's, as the best approximation extant for treating the dielectric properties of such liquids.

Fowler and Guggenheim's *Statistical Thermodynamics*[10] contains a valuable summary of dielectric theory mainly on the lines of Van Vleck[11] giving approval to Onsager. Since this is for most purposes a standard work, it appears necessary to answer in detail the arguments of such a reputable authority. The reader will be assumed to have access to the book. The "hindering energy" theory of Debye is not included, so that

---

[10] Fowler and Guggenheim, *Statistical Thermodynamics*, Cambridge, 1939.

[11] Van Vleck, *J. Chem. Physics*, 1937, **5**, 320, 556.

Onsager's theory is compared, rather unfairly, with a theory called Debye's, applied in a way which was never intended. There are two trivial misprints: the factor $(n^2 - 1)/(n^2 + 2)$ appears inverted in equation (1428, 3) and μ_0 is omitted from (1428, 9). Some caution is displayed with regard to Onsager's identification of the molecular radius a as the radius of a sphere of volume equal to the molecular volume at (1428, 16), but this is rather late to be cautious since the assumption was introduced at (1428, 1): it follows that (1428, 15) is in no way more accurate than Onsager's own formula (1428, 17). These are trivialities: a rather more serious point arises in paragraph (1429), where Onsager's result is expanded in a series and shown to correspond in the first three terms with a series obtained by Van Vleck by an entirely different argument. The convergence conditions for Van Vleck's series are not known. Those for the expansion of Onsager's result require the dielectric constant to be less than 1·85 (from 1429, 2 and 1429, 4). This is not the case for any polar liquid and certainly not for any liquid satisfying the opening condition $(n^2 - 1)/(\delta - 1) \ll 1$. ($\delta$ is the dielectric constant and n refractive index). If we take a compressed gas to satisfy the conditions the validity of Onsager's molecular model becomes more dubious than ever. (Either n must be taken as the refractive index of the "molecular substance" in which case the opening condition can never be satisfied or it is to be taken as that of the compressed gas, in which case the assumption of uniform refractive index over a cell equal to molecular volume is very far from reality.) It may also be remarked that the agreement found depends on the use of the "continuum value' for Q (*i.e.* 35·1). In a liquid one might plausibly prefer the value for spherical close packing (16·8) or one of the values for other packings of lower co-ordination (about 14) which make the term in $(\theta/T)^2$ small and *negative*. The comparison between Van Vleck's formula and "Debye's formula" (without hindering energy) is irrelevant. If Van Vleck's theory is correct, it corresponds to nature and hence is to be compared with Debye's formula *with* hindering energy which purports to do the same. Accepting Debye's theory, Van Vleck's formula gives us a theoretical value for the hindering energy at low density, namely if Debye's hindering field is defined as $-E/2\cos\theta$,

$$E = 2\sqrt{Q}\mu_0^2/V$$

though, owing to the uncertain convergence range of Van Vleck's formula, we cannot be sure this is valid for any concentration at which the effect is measurable.

There is one more general argument which must be taken up. In paragraph (1427), "Inadequacy of Debye's formula," it is indicated that any theory which permits the dielectric constant to pass to infinity below a certain temperature is in fundamental disagreement with reality: this is taken as one of the chief reasons for preferring the Onsager theory. It is true that this does not happen in liquids. It occurs for example in crystals such as Rochelle salt. But the main point is that passage of the dielectric constant to infinity signifies that the molecules become oriented without any applied field, that is to say the liquid crystallises: this occurs quite commonly in nature. There are many departures from the most elementary theory. The temperature of rotational freezing is only given extremely roughly. The transition is usually abrupt, and passes directly to the solid rather than to a liquid crystal. The stable crystal form produced is more often non-polar than polar, and when polar it is only exceptionally reversible by an applied field as in Rochelle salt. All these facts find ready explanation, and that related to the concepts which belong to Debye's hindering energy theory.

The present writer, therefore, remains an unrepentant protagonist of Debye's theory. If for quantitative calculation it is unreliable, it is still the best we have, while for qualitative interpretation it has no rival.

Summary.

Onsager's and Debye's theories of the dielectric constant of condensed polar media are critically compared. It is concluded that Debye's theory may be inaccurate but Onsager's certainly is. Arguments indicating that Debye's theory is very far wrong are refuted.

The author cannot conclude without expressing regret at contradicting Fowler and Guggenheim, to both of whom he owes a debt of gratitude. Although this note was written recently, those points in it which concern their book were put to the late Professor R. H. Fowler in 1940. A short discussion with Professor N. F. Mott has helped the author to eliminate certain errors from an earlier draft, but he alone takes responsibility for controversial views expressed.

H. H. Wills Physical Laboratory,
Royal Fort,
Bristol, 8.

SOME MATHEMATICAL MODELS REPRESENTING POLAR MOLECULES IN CRYSTALS

By F. C. Frank.

Received 8th April, 1946.

This paper contains an abstract of the results of a series of calculations made before the war and left unpublished because of it. Fuller publication may be made later.

In 1930 Pauling [1] was led by some considerations about solid hydrogen to the conclusion (not entirely new) that molecules of low moment of inertia might rotate freely in crystalline solids, generally with a transition at some temperature below the melting-point to a non-rotating state; and to the successful prediction that such solids, above the transition point, would have high dielectric constants (similar to the liquid) if the molecules were polar. That the same occurs with molecules of large moment of inertia, such as camphor, that there are often multiple transitions, with various changes of dielectric constant, and the evidence from dielectric dispersion and absorption,[2] show that Pauling's original theory was excessively simplified—as was indeed undenied from the start. Debye had already shown [3] that a high dielectric constant will result, without free rotation, if the polar molecule is able occasionally to change orientation between two or more alternative positions. Eucken [4] showed on thermal evidence that rotation was not necessarily free in the so-called rotating states of crystals. Nevertheless the idea persisted that a high dielectric constant signifies free rotation (for example, in 1940 Van Vleck [5] uses the earlier Pauling conception as a cardinal point in arguing against Debye's concept of hindered rotation in liquids). Consequently the present writer undertook to investigate a range of simplified models to show the effect upon dielectric constant, and its change with temperature, of various degrees of restriction upon molecular orientation, with or without rotation, which could be expected to exhibit the principal characteristics of restrictions occurring in actual crystals. It may be remarked in passing that there is no theory of dielectrics which does more than set up models which can be hoped to exhibit properties illustrative of actuality.

---

[1] Pauling, *Physic. Rev.*, 1930, 36, 430.
[2] Frank, *Trans. Faraday Soc.*, 1936, 32, 1634.
[3] Debye, *Polar Molecules*, 1929, p. 104.
[4] Eucken, *Z. Elektrochem.*, 1939, 45, 125.
[5] Van Vleck, *Ann. New York Acad. Sci.*, 1940, 40, 293.

Although the basic ideas are related to Debye's theory of hindered rotation,[6] the calculations are conducted without essential regard to the vexed question of the way in which the inner field, acting on the molecule, depends on the field applied for measurement from without; whether according to Mosotti's rule or to some other. (The author actually considers Mosotti's rule the best approximation known, at least for crystals.)

The method adopted is the simple "classical" one, as used by Debye, in which, firstly, the Boltzmann distribution is stated for the allowed orientations of the molecule, subject to the restrictions intended to represent the effect of intermolecular forces in the crystal, which define each model investigated. The weighted mean resolved dipole moment of this distribution is $\bar{m}$ (not exceeding the moment, μ, of the molecule). As the next step the applied measuring field is supposed to produce a field F at the molecule, and the resulting vector change in mean resolved moment, $\delta\bar{m}$, due to the altered distribution, is investigated. Its mean over-all orientations of the applied (weak) field, $\overline{\delta\bar{m}}$, is proportional to F and collinear with it. As a final step, the proportionality factor (polarisability) $\alpha = \overline{\delta\bar{m}}/F$ is compared with the value for free dipoles, giving us the reduction factor $R = \alpha/(\mu^2/3kT)$. In the course of the calculation α is evaluated for fields parallel to the three principal axes, relative to the hindering field, but only the mean value is quoted in this abstract. Unless the restrictions are in some way asymmetric, they have no effect on the polarisability: *viz.*, if $\bar{m} = 0$, $R = 1$. On the other hand, if there is an asymmetry (*e.g.* in potential energy) it is found that the effect of additional, symmetrical, restrictions is important.

Results.

(I) Asymmetry in Potential Energy Only.

In all of the following set of models, the highest and lowest potential energies of permitted orientations of a molecule differ in energy by E. Unless otherwise stated, these correspond to opposed orientations, so that the *asymmetrical component* of the hindering field takes the form $V = -\frac{1}{2}E\cos\theta$. We use the abbreviation $\gamma = E/2kT$. In this first set of models, permitted positions have equal statistical weights.

Two opposed positions.—This is the simplest possible case, and yields the result

$$\bar{m}/\mu = \tanh\gamma$$
$$R = \operatorname{sech}^2\gamma.$$

It is included in the case examined by White[7] for two positions differing in potential energy by E, making an angle β with each other, for which

$$R = (\sin^2\beta/2)\operatorname{sech}^2\gamma.$$

Four mutually perpendicular positions in a plane.—Two opposed positions differ in potential energy by respectively $+E/2$ and $-E/2$ from two other positions perpendicular to these in the same plane.

$$\bar{m}/\mu = \tanh(\gamma/2)$$
$$R = \operatorname{sech}^2(\gamma/2).$$

Six mutually perpendicular positions.—Of the six positions, two opposed positions differ in energy by respectively $+E/2$ and $-E/2$ from the rest.

$$\bar{m}/\mu = (\sinh\gamma)/(2 + \cosh\gamma)$$
$$R = (5 + 4\cosh\gamma)/(2 + \cosh\gamma)^2.$$

---

[6] Debye, *Physik. Z.*, 1935, **36**, 100, 193.

[7] White, *J. Chem. Physics.*, 1939, **7**, 58.

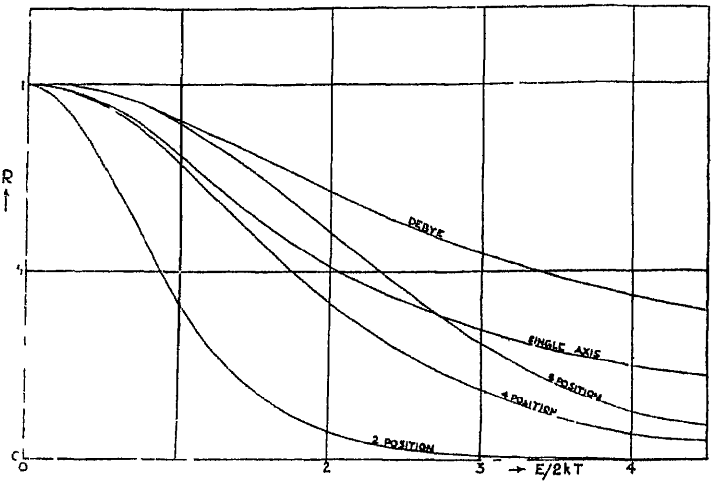


FIG. 1 (1).—Variation of polarisation reduction factor, with hindering energy or reciprocal temperature for full rotator (Debye), single-axis rotator, and the 2-, 4-, and 6-position cases.

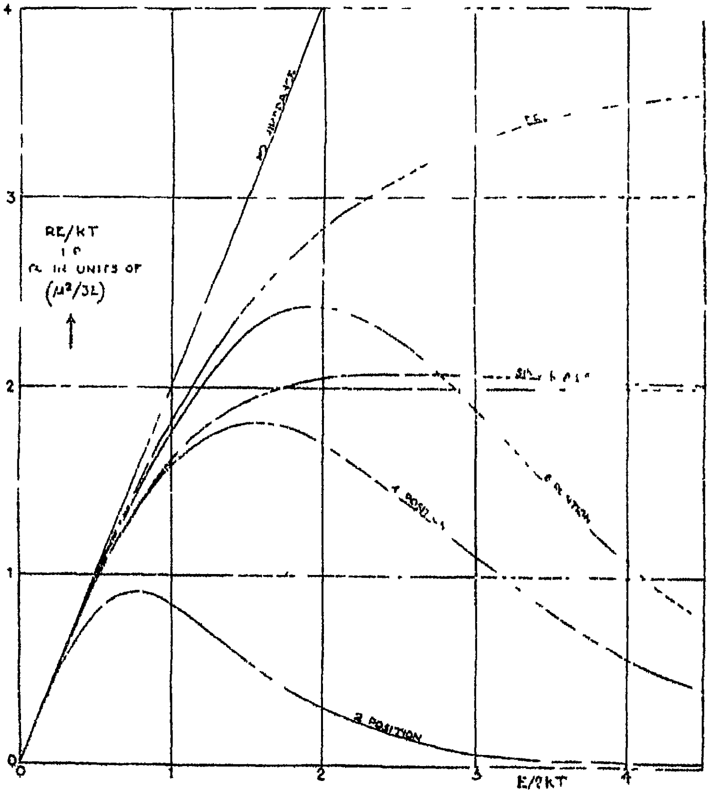


FIG. 1 (2).—Variation of polarisability, for the same cases as Fig. 1, with reciprocal temperature, for constant hindering energy.

Rotation about a single axis.—The dipole lies in a plane, subject to the potential field

$$V = -\tfrac{1}{2}E\cos\theta$$

$$\bar{m}/\mu = I_1(\gamma)/I_0(\gamma)$$
$$R = 1 - I_1^2(\gamma)/I_0^2(\gamma).$$

(Here we employ the notation of the British Association Tables of Bessel functions, where $I_n(x)$ signifies $(-\iota)^n J_n(ix)$.)

Full rotation, subject to the potential $V = -\tfrac{1}{2}E\cos\theta$. This is Debye's model for hindered rotation[6] (though Debye defined the potential as $V = -E\cos\theta$: *i.e.* our E is twice as great as his).

$$\bar{m}/\mu = \coth\gamma - 1/\gamma$$
$$R = 1 - (\coth\gamma - 1/\gamma)^2.$$

For all these cases,

$$R = 1 - (\bar{m}/\mu)^2.$$

These results are exhibited in Figs. 1(1) and 1(2). Some of them appear again in Fig. 3 (1). An important point to notice is that when γ is not too large, the 4-position model closely resembles the single-axis rotator, and the 6-position model the full rotator. On the other hand, for large γ the discrete position models show a regime in which α decreases with fall of temperature, the fall being fairly rapid whenever $R > 0{\cdot}5$. Further explanations are needed for small values of R if α continues to rise with falling temperature and dispersion at moderate frequencies indicates the need for a discrete position model.

No single substance can be expected to show the behaviour represented by the curves of Fig. 1 (2) over a very wide range, unless the dipoles form only a small part of the total material, for otherwise the changing degree of orientation must be expected to alter E, and crystal transformations will commonly supervene. We shall return to this point.

(II) Models with Asymmetry in Energy and Entropy.

The models of class (I) clearly lack generality, since in general we must expect the sundry possible orientations of a molecule to differ in libration frequency and consequently in entropy or statistical weight.

Two opposed positions differing in energy and entropy.—One of the two positions has not only a potential energy greater than that of the other by E, but also its statistical weight is g times that of the position of lower energy. The entropy difference is $k\,.\ln g$.

$$\bar{m}/\mu = (e^{\gamma} - e^{-\gamma})/(e^{\gamma} + ge^{-\gamma})$$
$$R = \operatorname{sech}^2(\gamma - \tfrac{1}{2}\ln g) = 1 - (\bar{m}/\mu)^2.$$

This is exhibited in Figs. 2 (1) and 2 (2). If $g > 1$, we can have a range of temperatures over which the polarisation, though reduced below the value for free dipoles rises more rapidly with falling temperature. Thus from the temperature coefficient an abnormally high dipole moment would be found. This sort of anomalous result is known in solid hydrogen iodide.[8] This model can also provide small values of R while the polarisability still rises with falling temperature.

Free rotator with one frozen position.—A rather different model which may be considered related to the same class is one which provides the simplest representation of the common concept for "rotational transitions". The molecule can rotate freely in all directions, or "freeze" into one particular direction: this position is assigned an arbitrary statistical weight, and a potential energy $-E$ relative to all other directions.

[8] Smyth and Hitchcock, *J. Amer. Chem. Soc.*, 1933, 55, 1830; 1934, 56, 1084.

If in the absence of an applied field there are n_r rotating molecules and n_f frozen molecules, where $n_r/n_f = ge^{-2\gamma}$, then

$$\bar{m}/\mu = n_f/(n_r + n_f) = 1/(1 + ge^{-2\gamma})$$
$$R = 1 - 1/(1 + ge^{-2\gamma})^2 = 1 - n_f^2/(n_r + n_f)^2.$$

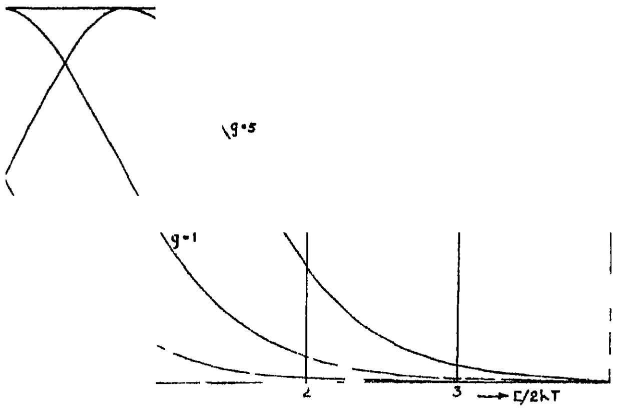


FIG. 2 (1).—Variation of polarisation reduction factor with hindering energy or reciprocal temperature, for 2-poisition case with differing statistical weights.

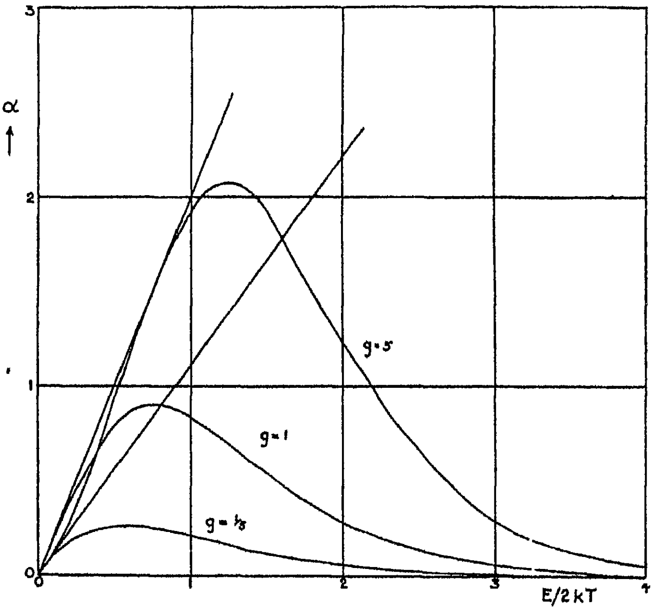


FIG. 2 (2).—Variation of polarisability with reciprocal temperature, for 2-position case with differing statistical weights.

It is to be noted that this value of R is not the same as the proportion of unfrozen dipoles, which is $\{1 - n_f/(n_r + n_f)\}$. When half are frozen, $R = 0{\cdot}75$ and when 10 % are frozen, $R = 0{\cdot}99$. The reason is that the applied field not only orients the freely rotating molecules, but also freezes in a few whose freezing position is pointing its way, and melts out some more from opposing positions.

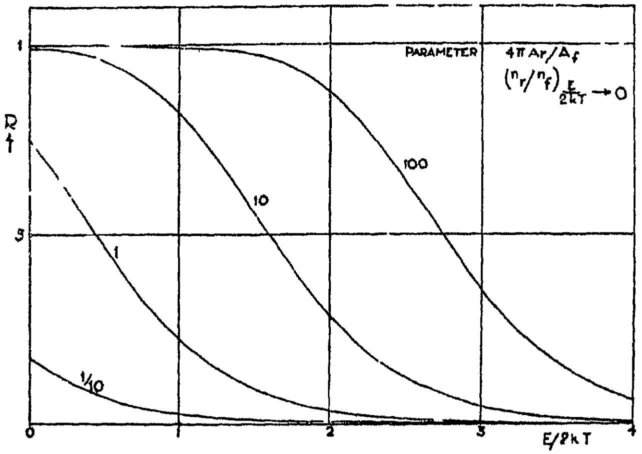


FIG. 2 (3).—Variation of polarisation reduction factor with hindering energy or reciprocal temperature for free rotator with one freezing position.

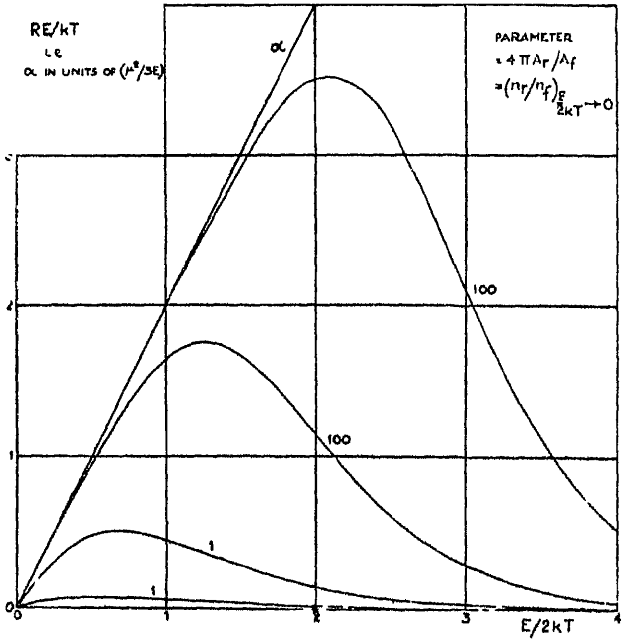


FIG. 2 (4).—Variation of polarisability with reciprocal temperature for free rotator with one freezing position.

(III) Co-operative Models.

In the foregoing models, the influence of the environment of a polar molecule has been represented by a fixed field. When the environment is one of similar polar molecules, no less free to move, this is an imperfect representation. We must therefore consider co-operative effects, and distinguish two sorts, namely the thermal co-operative effects, initiated by change of temperature, and electrical co-operative effects initiated by an applied electric field.

The Thermal Co-operative Effect: "Bragg-Williams"

The asymmetry in the molecular environment field must be, at least in part, of electrostatic nature—the resultant electrostatic field of all the surrounding dipoles. This can be pictured from two (mainly equivalent) viewpoints. In a crystal of polar molecules, from which one is removed, there will be in general a definite electrostatic field in its cell caused by the dipoles of all the rest; on replacing this molecule it would tend to be oriented by this field. Alternatively we may picture around every dipole, even in the liquid, a "dipole orientation atmosphere," analogous to the "ionic distribution atmosphere" in the Debye-Huckel theory of strong electrolytes (*cf.* H. Müller [9]). With rise of temperature and consequent diminution of order, this average field will usually decrease, and the simplest assumption is that it is proportional to the mean resolved moment, $\bar{m}$. (Similar calculations were made by Fowler.[10])

Our basic assumption for the co-operative effect is

$$E = E_0 \bar{m}/\mu.$$

Combining this with the result obtained in (1) for the *two-position case*:

$$\bar{m}/\mu = \tanh \gamma$$

we have

$$\gamma/\tanh \gamma = E_0/2kT$$

which defines an equilibrium value of γ for every temperature T. It can be verified that this is a stable equilibrium. As before

$$R = \operatorname{sech}^2 \gamma.$$

The result is a typical second order transition (Fig. 3 (1)). $\bar{m}/\mu$ falls from 1 at absolute zero to 0 at $T = T_c = E_0/2k$, while R changes from 0 to 1 in the same range. Above T_c, $\bar{m} = 0$ and $R = 1$. With the same basic assumption for the co-operative effect, we obtain similar second order transitions for the

Single axis rotator with $V = -\frac{1}{2}E \cos\theta$:

$$\gamma I_0(\gamma)/I_1(\gamma) = E_0/2kT$$

$$T_c = E_0/4k.$$

Full rotator with $V = -\frac{1}{2}E\cos\theta$ (Debye's model):

$$\gamma/(\coth\gamma - 1/\gamma) = E_0/2kT$$

$$T_c = E_0/6k.$$

Full rotator with $V = -\frac{1}{2}E_1\cos\theta - \frac{1}{2}E_2\cos 2\theta$.

This is a model which includes as extremes both Debye's model and the 2-position case.

With $\gamma_1 = E_1/2kT, \quad \gamma_2 = E_2/2kT$

$$\bar{m} \quad \frac{\int_{-1}^{+1} e^{\gamma_1 x + 2\gamma_2 x^2} x \, dx}{\int_{-1}^{+1} e^{\gamma_1 x + 2\gamma_2 x^2} dx}$$

With the co-operative condition $E_1 = E_0\bar{m}/\mu$, we find

$$T_c = \{0{\cdot}3333 + 0{\cdot}0889(2\gamma_2) + 0{\cdot}00847(2\gamma_2)^2 - 0{\cdot}00113(2\gamma_2)^3 - 0{\cdot}000344(2\gamma_2)^4 + 0{\cdot}000049(2\gamma_2)^5 + \ldots\}E_0/2k.$$

---

[9] Müller, *Physik. Z.*, 1937, 38, 501.

[10] Fowler, *Proc. Roy. Soc. A.*, 1935, 149, 1; 1935, 151, 1.

At very low temperatures this reduces to a harmonic oscillator, with polarisability

$$\alpha_0 = 4\mu^2/(3E_1 + 12E_2).$$

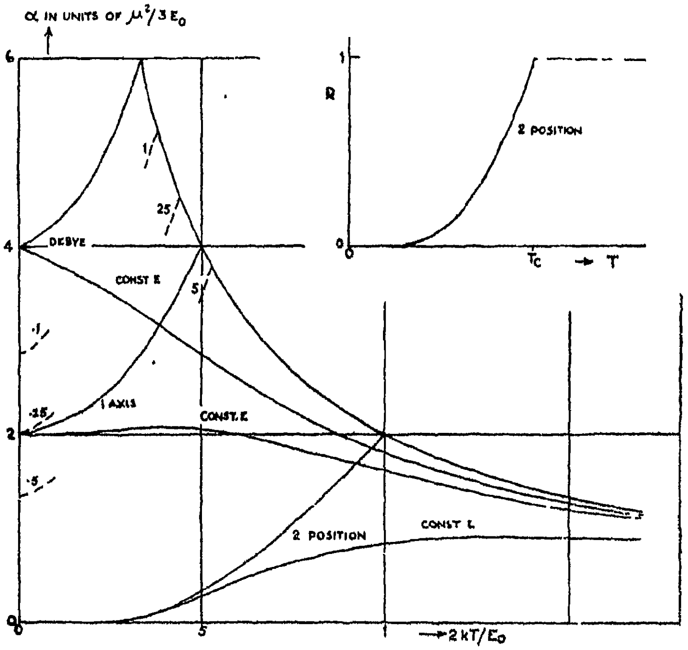


FIG. 3 (I).—2nd order transitions. Variation of polarisability with temperature when hindering energy varies "co-operatively" ($E = E_0\overline{m}/\mu$), for full rotator (Debye), single axis rotator, 2-position model, and three cases of Debye's model modified by a cos 2θ term. Cases with constant $E = E_0$ are shown for comparison. *Inset.*—Polarisation reduction factor-temperature for 2-position model.

If E_2 is allowed to vary co-operatively in proportion to $\overline{m}$, T_c becomes $E_0/6k$ for all cases.

TABLE I

	Full Rotator. $V = -\frac{1}{2}E_1 \cos\theta - \frac{1}{2}E_2 \cos 2\theta$ $E_1 = E_0\overline{m}/\mu = E_0$ at $0°$ K.					*Single Axis.* $V = -\frac{1}{2}E\cos\theta$ $E = E_0\overline{m}/\mu$.
E_2/E_0	0 (Debye model)	0·1	0·25	0·5	∞ (2-position)	0
$T_c/(E_0/2k)$	1/3	0·382	0·442	0·528	1	1/2
$\alpha/(\mu^2/3E_0)$ at $T = T_c$	6	5·23	4·52	3·79	2	4
at $T = 0$	4	2·858	2	1·333	0	2

This is a sufficient number of examples to demonstrate that for all models of class (I), with asymmetry in potential energy only, a simple "Bragg-Williams" type of co-operative condition leads only to second order transitions. For a given energy-asymmetry the addition of symmetrical restrictions raises the transition temperature, and the temperature at which the "co-operative rise" becomes appreciable (Fig. 3 (1)), but in principle the curves are all of the same type. This is satisfactory, since this type of orientational transition is quite common, as are others approximating to it in the sense of having a large pre-transition range though terminating with a small first order discontinuity. However, there are other orientational transitions which are unambiguously of first order, having relatively little pre-transition, and we must look further for models exhibiting this behaviour.

Co-operative Changes Producing First Order Transitions.

Such models are provided by applying the same co-operative condition to the energy-asymmetry for models of class (II), *i.e.* those having asymmetry in statistical weight. Let us consider in detail the

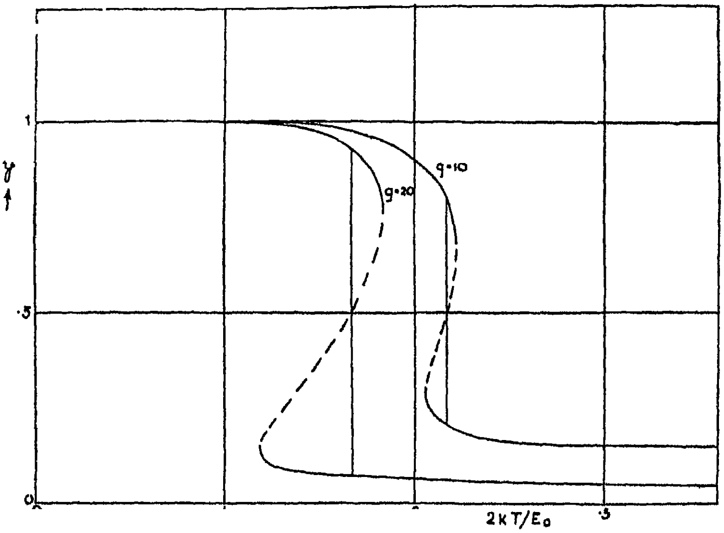


FIG. 3 (2).—Variation with temperature of number of frozen molecules for the full rotator with one freezing position, co-operative model.

Free rotator with one frozen position. The simple electrostatic interpretation of E is no longer appropriate, but a "Bragg-Williams" type of co-operative assumption is still reasonable. We have

$$n_r/n_f = g\mathrm{e}^{-E/kT}$$

and assume for the co-operative condition

$$E = yE_0 \text{ where } y = n_f/(n_r + n_f) = \overline{m}/\mu.$$

By combining these

$$T = E_0 y/k \ln\left(\frac{gy}{1-y}\right)$$

If $\ln g$ exceeds 2, there is now a retrograde portion in the graph of y against T (Fig. 3 (2)). This will be the normal case, for $R \ln g$ is the entropy difference for molecules rotating freely or librating about the frozen position, which has been estimated[11] for a number of diatomic and triatomic molecules, ranging between 4·85 and 7·6 cal./mole/degree. The corresponding values of g range from 11 to 45. In the retrograde range there are three equilibrium values of y for each temperature, the middle one being an unstable equilibrium. Of the other two stable equilibria, one must be more stable (*i.e.* have lower free energy) than the other, except at one transition temperature, which can be shown by statistical-thermodynamic reasoning to occur when $y = 1/2$ for the unstable equilibrium. This transition temperature is

---

[11] Hirschfelder, Stevenson and Eyring, *J. Chem. Physics*, 1937, 5, 896.

$$T_t = E_0/2k \ln g.$$

The orientational heat content, and the polarisation reduction factor R are both given by the same curve (Fig. 3 (3)) being both proportional to $(1 - y^2)$.

The curves obtained reproduce the essential characteristics of known first order orientational transitions, *e.g.* in solid HCl [8, 12] or the long chain ketones.[13] There is a large abrupt change of polarisation at the transition temperature, but a significant continuous change below it. Above the transition, about 90 % of the polarisation is due to molecules free to rotate, and for the rest the potential energy minimum is shallow, so that any possible dispersion will occur at very high frequencies. Below the transition, half the polarisation is due to change of orientational distribution of the librating molecules, the field lifting some of them out of their potential minima, and driving others in. Here dispersion can be expected and is observed.[12] It is also observed, in agreement with our result, that this dispersion affects only a part of the remaining temperature-dependent polarisation below the transition. The model is of course too simple, for both dielectric and thermal data [4] indicate that in HCl the rotation is still somewhat "hindered" above the transition point.

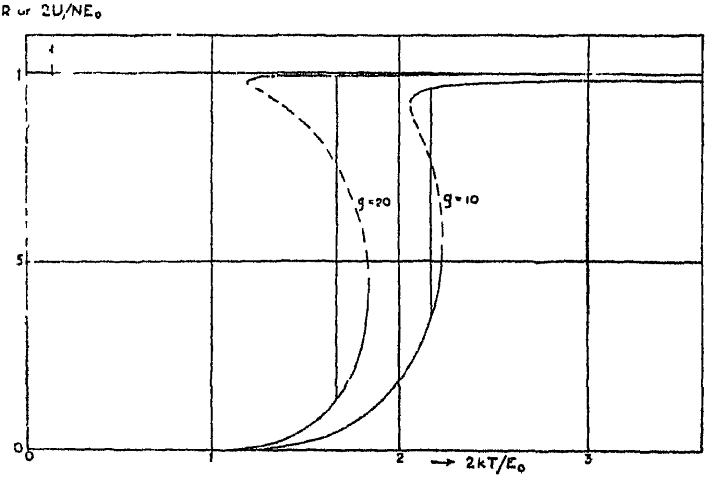


FIG. 3 (3).—1st order transition: Variation with temperature of polarisability reduction factor, or of heat content, for the full rotator with one freezing position, co-operative model.

The two position model, treated in the same way, gives first order transitions as soon as g exceeds 1 (*i.e.* whenever the state of higher energy has the lower vibration frequency).

It is to be noted that first order transitions are only obtained in this way if the entropy difference persists while the energy difference diminishes. This signifies that libration frequencies remain relatively unaltered while the depth of the potential minimum (or the difference in depth between two minima) changes.

---

[12] Hettner (G. and E.) and Pohlman, *Z. Physik.*, 1937, 108, 45.
[13] Muller, *Proc. Roy. Soc.*, *A*, 1937, 158, 403.

(IV) Electrical Co-operative Effects.

There is a mode of dipole-dipole interaction not allowed for (except partially and imperfectly if one uses the Mosotti formula) in any of our previous models. Unlike the interactions considered heretofore, it can lead to an enhancement of the orientation polarisation. When enhanced polarisations occur in dilute liquid solutions, it is usually supposed that pairs of molecules adhere as (*a*) and wander freely into random orientations (*b*) (Fig. 4 (1)). The polarisation is increased from $n\mu^2$ to $\frac{1}{2}n(2\mu)^2$, *i.e.* it is doubled. This cannot happen in solids. However, if in such a case as a solid solution of an ester or a ketone in paraffin wax there is a tendency for dipoles to be grouped in pairs, and by rotation of the molecules individually they can assume the alternative positions (*c*) and (*d*), the result will be just the same doubling of the polarisation. These are, of course, the two equilibrium positions determined by the dipole interaction in such pairs, and are the only positions that significantly occur if the dipole interaction is strong (*i.e.* if $\lambda = \mu^2/kTr^3 \gg 1$, where r is the distance between the dipoles). Actually, λ is just about 1 for typical cases. Let us consider the case of a pair of coplanar dipoles, free to rotate about parallel axes perpendicular to their plane, and interacting only by their electrostatic fields.

$$V = -\frac{\mu^2}{r^3}(2\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2).$$

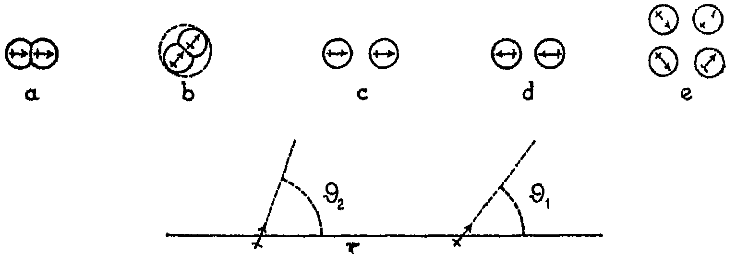


FIG. 4 (1).

The mean polarisability for all directions of applied field is

$$\alpha = \frac{2\mu^2}{3kT}\left(1 + \frac{1}{4}\lambda - \frac{1}{128}\lambda^3 + \frac{1}{3072}\lambda^5 - \ldots\right)$$

Thus for $\mu = 2D$, $T = 300°$ K., and $r = 4{\cdot}5$ A., $\lambda = 1{\cdot}06$ and the polarisability is enhanced by 26 %. With stronger interaction the enhancement is increased to a limit of 100 %, or to higher limits for larger groups of molecules. The groups need not be linear. There will also be an enhancement for such groupings as (e), and for all polar groupings, such as occur in piezo-electric and pyro-electric crystals, provided only that for n molecules the moment exceeds $\mu\sqrt{n}$, and the group is capable of reversing.

When the group becomes so large as to constitute a macroscopic crystal, and there is still a reversal mechanism, we arrive at a body showing the "ferro-electric" properties of Rochelle salt and the acid phosphates. In most polar crystals, the ability to reverse is absent: when it occurs, it must be because a small reversed region is stable, and a reversal boundary labile, so that if the reversed region is favoured by a field, it can grow.

The "ferro-electric" substances provide the most striking example of such phenomena, but they are probably of general significance, for liquids as well as solids. The structure of a liquid, like that of a solid, is ruled mainly by short distance forces, which favour certain spacial relationships between molecules, and it is unthinkable that they should totally succeed in maintaining these relationships just below the melting-point and totally fail just above it. All evidence is to the contrary, in both respects. We have to suppose that a substance in the liquid state constantly sketches within itself, here, there and everywhere, in local groupings, albeit without reaching perfection or great expanse, all the crystalline structures of which it is capable, as well as other equally orderly but non-crystallographic groupings. These latter are orderly close-formed groups of small numbers of molecules which do not have the right symmetry properties for infinite regular extension. They are doubtless important in all cases and dominant in glass-forming liquids. Apart from them, the most favoured groupings should in general be those which are actually realisable as crystal forms, being the stablest, but with a preference for those of higher entropy at the cost of a somewhat higher energy, these being the high temperature forms, sometimes metastable or unrealised.

The same considerations apply to the orientational order in a locationally crystalline body.[14] In either case such states of order of limited extent are termed quasi-crystalline. The effect on the dielectric constant of the quasi-crystalline structure depends on the type of groupings present, the electrically co-operative tendency will occur if there is a substantial proportion of polar groupings. When such polar groupings occur, an applied field will prolong the life of those pointing in its own direction, at the expense of those pointing the other way. A superposed thermal co-operative effect can dramatically amplify the phenomenon: it has been demonstrated by Jaffé[15] that at a second order transition between polar and non-polar crystal forms the dielectric constant must pass to infinity, as it does in Rochelle salt or HBr. When polar groupings are rare, then the electrical co-operative tendency falls away, since a small applied field does not favour any one non-polar crystal form above another In such cases the " hindering energy " treatment is applicable.

The type of grouping formed depends on molecular shape, on the locational lattice structure, and on specific chemical properties. Polar formations are favoured by molecules with greatest extension perpendicular to the dipole, by crystal structures containing lines of polar radicals, and by the special associative mechanism of the hydroxyl bond (particularly in solutions at intermediate dilution, where solute molecules can associate in groups of small numbers). Such considerations indicate that it is hopeless to expect final accuracy from any " sphere-and-continuum " theory of dielectrics.

Summary.

Various restrictions on molecular orientation in a crystal, both with and without rotation, and having asymmetry in potential energy or statistical weight, are considered and the resulting dependence of dielectric polarisability on temperature ascertained. When thermal co-operative effects are introduced, the models lead to second order and first order transitions, such as are observed. Electrical co-operative effects are discussed.

H. H. Wills Physical Laboratory,
Royal Fort,
Bristol, 8.

---

[14] Frank and Wirtz, *Naturwiss.*, 1938, **26**, 697.
[15] Jaffé, *Physical Rev.*, 1938, **53**, 917; 1937, **51**, 43.

GENERAL DISCUSSION

Mr. D. M. Ritson (*Oxford*) said: I have read with great interest Kirkwood's paper on the Dielectric Constant. There is, however, one small point which I should like to raise. In his derivation of his statistical mechanical formula, too little attention appears to have been paid to the effect of the induced polarisability. The following considerations would seem to indicate that slight modification is necessary.

In equation (5), Kirkwood puts the increase of potential energy of a fixed configuration of dipoles occupying a spherical region of space with total moment M, on applying a field E_0 as equal to $- E_0M$. The terms for the effect of the induced polarisation are, he says, of the order of E_0^2 and can accordingly be ignored.

If we can consider the problem from the viewpoint adopted by Böttcher, we should identify the internal field with Böttcher's $\vec{E}_r$. If the dielectric constant due to the atomic and electronic polarisability is ϵ_0 and N is the number of molecules per cc., we should arrive at the following value for $\vec{E}_r$.

$$\vec{E}_r = \frac{12\pi\epsilon_0 + \frac{2}{Nr^3}(\epsilon_0 - 1)^2}{(2\epsilon_0 + 1) \cdot 4\pi} E.$$

The field E is the macroscopic electric field inside the specimen and is related to E_0 as follows:

$$E = \frac{3}{\epsilon_0 + 2} \cdot E_0.$$

thus giving

$$\vec{E}_r = \frac{3}{4\pi} \cdot \frac{12\pi\epsilon_0 + \frac{2}{Nr^3}(\epsilon_0 - 1)^2}{(\epsilon_0 + 2)(2\epsilon_0 + 1)} \cdot E_0.$$

Thus we see that the potential energy would not be increased by an amount $- E_0M$, but by the following amount

$$\frac{3}{4\pi} \cdot \frac{12\pi\epsilon_0 + \frac{2}{Nr^3}(\epsilon_0 - 1)^2}{(2\epsilon_0 + 1)(\epsilon_0 + 2)} \cdot E_0M.$$

Only if the value of $1/Nr^3 = 4\pi/3$ (Onsager's assumption) does the expression reduce to that obtained by Kirkwood.

Dr. F. C. Frank (*London*) said: We are still far from a perfect theory of the liquid state, or even of the simpler case of a locationally crystalline, but orientationally only quasi-crystalline solid. Consequently, in all our theories of dielectric constant, we are bound to approximate at some stage. There are various ways of approach. The classical physicist will take some such model as a spherical molecule, with an ideal dipole at its centre, situated in a structureless medium, and thereafter apply straightforward accurate mathematics. This is Onsager's method. The chemist knows there is no such molecule or medium in nature: he will probably take a molecular model corresponding much more closely to reality, and treat it with somewhat crude mathematics. I have used this method myself.[1] Debye's theory of dielectric constants in polar liquids inclines towards the same class, though his mathematics should not be called crude, but slightly happy-go-lucky, since he had only his (often very reliable) intuition to justify his use of the Mosotti field well outside its proven range of validity. The modern physicist's approach is to start from very general principles, and rigorous mathematics, with which he produces some very elegant theorems, but seldom a complete result which can be applied directly to any but the very simplest of actual substances. To get that it is necessary to make pretty broad mathematical approximations or to introduce rather arbitrary physical models. The merit of the method is that when the basic theory has been worked out, it may be quite easy to try many models in turn. Such methods are familiar in the quantum mechanics of molecules : they are also the methods of Van Vleck and Kirkwood in this problem. Van Vleck, when he came to the approximation and model-making stage, adhered too closely to the unreal model of Onsager, believing at first that he had found independent justification of it. Though he subsequently receded from this position,[3] he still rejected the " hindering energy " theory because he attached too much faith to the very simplified theory of molecular rotation in solids, produced by Pauling in 1930. I am very pleased to see that Kirkwood now incorporates hindering energies when he reaches the model-making stage in his accurate theory, and arrives at results intermediate between those of Onsager and Debye. There is good hope that on these lines we shall progress to a more precise theory of dielectric constants. In the meantime we have the Debye theory, incorporating various developments of the idea of hindering energy, as a good general approximation which is never very far wrong. In particular the development of this theory incorporating thermal and electrical co-operative effects, as originated by Fowler, and rendered doubtful by the work of Onsager, can be reinstated.

---

[1] Frank, *Proc. Roy. Soc.*, *A*, 1935, 152, 171.

I am very happy to find that Professor Bauer's account relieves me of the need to say much in explanation of my second paper. Much of my paper is concerned with calculating particular cases according to almost exactly the same ideas as have been treated with greater generality by Professor Bauer. I am only sorry that circumstances prevented us both from publishing and discussing our results much sooner.

Prof. E. Bauer (*Paris*) said : I have also tried to develop a dielectric theory analogous to Weiss's theory of ferromagnetism, in order to explain what happens near transition points of second order. But we must take into account an important fact : the relaxation time increases very much near these transition points, and then suddenly falls to a very small value above this temperature.

Dr. F. C. Frank (*London*) (in reply to Professor Bauer) said : According to Peierls, and Landau, there must always be a discontinuity, of some order, when there is a change of macroscopic symmetry at the transition. But in the case that the disordered molecules (or ions) are situated in an ordered crystalline matrix, their order may remain crystalline—not quasi-crystalline —even when it is almost nil, because the persisting structure can define orderly positions for them. Then the change from almost perfect order, and low dielectric constant, to almost perfect disorder, and high dielectric constant, can take place without necessarily showing any discontinuity.

(*Note added subsequently* : However, on reference to the curves shown for an alum in Professor Bauer's original paper, it is seen that this is not a good case of the phenomenon described. The curves shown are temperature-dispersion curves down to the lowest frequency used (50 c./s.). If such a continuous change in the static dielectric constant does occur in this case, it must be at still lower temperatures, and hidden from us by the dispersion, unless we can use much lower frequencies.)

According to Jaffé, the dielectric constant must go to infinity at any second order transition between two crystal forms, one of which has a polar axis and the other has not. In simple terms, at the transition point the applied field can enforce the transition with infinite ease, because there is no heat of transformation. Actually, all such transitions (and indeed all transitions in solids) show hysteresis, and a range of co-existence of both forms, owing to the presence or development of stresses. Then the process of dielectric polarisation is a crystal growth phenomenon, and it is suggested that this can account for a great lengthening of relaxation time.

Prof. E. Bauer (*Paris*) (*communicated*) : If, without changing the shape of the crystalline field, one makes it proportional to an order parameter

---

[3] Van Vleck, *Ann. New York Acad. Sci.*, 1940, 40, 293.

(say, to $1 - x$, x being the fraction of molecules which rotate freely), the relaxation time would fall continuously near the transition point. This is never the case. It is necessary to make a precise discussion of experimental facts.

I have tried to do this for HBr and NH_4I. This discussion led to interesting results concerning the change of shape of the potential barrier before it disappears at the transition point. Unfortunately the experimental data are not yet sufficient to bring this discussion to definitive conclusions. It would be very interesting to extend the measurements of Smyth and Hitchcock on HBr to all available wave-lengths of the Hertzian spectrum.

In reply to the above remarks of Dr. Frank, I do not think that the process of dielectric polarisation near a second order transition point is a crystal growth phenomenon, but an internal and molecular phenomenon.

According to experiments by Weighe and others, I think that in the present state of our knowledge we may consider two kinds of transitions :

(1) The *real* second order transitions, with no change of lattice-structure, no apparition nor disparition of any polar axis, and generally, no hysteresis. The transition point of HBr, where large dielectric anomalies are observed, seems to be such a point.

(2) The *apparent* second order transitions, with a small change of lattice structure. Those are in fact first order transitions, where the smallness of the discontinuity and the inner stresses allow the coexistence of two different crystalline phases during some non-vanishing interval of temperature and causes thermal hysteresis phenomena which have nothing to do with relaxation.

Anyway, all those questions are still rather obscure and much experimental and theoretical work has to be done before we can really understand the mechanism of all those phenomena.

Dr. C. J. F. Böttcher (*Rotterdam*) (*communicated*) : Dr. Frank has a serious objection to Onsager's method of calculation, in that a constant ϵ is taken for the surroundings of the molecule although it might be expected that in the interaction of a molecule with its neighbours dielectric saturation should occur. Dr. Frank calculates that the directing influence of two dipoles of 2 D. at a distance of 4 A. shows already saturation as the field of one dipole at the place of the other is about 10^7 v./cm. and hence $\mu F > kT$. Now I calculated [1] that the reaction field is 10^7 v./cm. for chloroform and $4 \cdot 2$ 10^7 v./cm. for nitrobenzene. So the total reaction field caused by all the neighbouring molecules equals approximately the field caused by one molecule in the special position considered by Dr. Frank. This means that such positions are so rare that they will hardly influence the calculation. It is my opinion that it is much more difficult to determine under which circumstances saturation in the interaction of the dipoles will occur than it is in the case of a dipole in a homogeneous directing field. But there is striking evidence for the fact that a saturation as meant by Dr. Frank does not occur.

1. In chemical language Dr. Frank's assertion should say that almost all polar liquids were associated whereas in fact only very few liquids are really associated (water, alcohols, acids).

2. Formula (10) of my contribution to this discussion can be written as :

$$\vec{E}_r = \frac{3kT}{3kT + \mu R} \cdot \vec{F} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Now for a dipole liquid with one component the induced moment P per cm.^{3} is given by

$$\frac{\epsilon - 1}{4\pi} \cdot E = P = N\left(\alpha F + \frac{\mu^2}{3kT} \cdot E_r\right) \quad . \quad . \quad . \quad . \quad (2)$$

Substituting $E \cdot (\epsilon + 2)/3$ for F and using (1) we get :

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{4}{3}\pi N\left(\alpha + \frac{\mu^2}{3kT + \mu R}\right) \quad . \quad . \quad . \quad . \quad (3)$$

So the mere introduction of the fact that the reaction field does not contribute to the directing field leads to the formula of van Arkel and Snoek[3] which is based upon the experimental data and this confirms that in first approximation the reaction field accounts for the deviations of Debye's formula. Thus it is to be expected that the small differences still existing between theory and experiment are only caused by the simplicity of the model used (spherical cavity, dipole in the centrum, etc.).

Dr. Fröhlich claims that Onsager's theory should hold only for $\epsilon > 9$ so he is less vague in his objections than Dr. Frank. If this were true it ought to be confirmed by the experimental data. Now from Onsager's theory the following formula can be obtained;[1]

$$\mu^2 = \frac{9kT}{4\pi N} \cdot \frac{(\epsilon - n^2)(2\epsilon + n^2)}{\epsilon(n^2 + 2)^2} \quad . \quad . \quad . \quad (4)$$

With the aid of this formula I calculated the dipole moment of a large number of organic[4] and inorganic[5] compounds over widely varying temperature ranges. The values obtained are in most cases in better agreement with the vapour values than those obtained from dilute solutions with Debye's formula. Anyhow there is no indication that there is a certain critical value ϵ_0 for ϵ in that for $\epsilon < \epsilon_0$ the formula holds better than for $\epsilon > \epsilon_0$. From my calculations it is quite sure that if such a critical value should exist it must be higher than 40 for I used more than a hundred measurements with $9 < \epsilon < 40$ for numerous compounds. It is even probable that the deviations occurring with "associating" liquids (water, alcohols, acids) are mainly caused by the eccentric place of the dipole as this association even occurs for low dielectric constants (acids).

I should like to ask Dr. Fröhlich what experimental evidence he has for his assertion that Onsager's theory should hold better for $\epsilon < 9$ or another critical ϵ value than for $\epsilon > 9$.

Finally I object to Dr. Frank saying that the oldest problem, concerned with the dielectric constant or refractive index of a medium composed of non-polar molecules can now be regarded as settled as he certainly means that the formulæ of Lorentz-Lorenz and Clausius-Mosotti are the solutions of these problems.

It is my opinion that the small deviations existing between the experimental data and these formulæ are very important and that the study of these deviations leads to many new perspectives as mentioned in the last paragraphs of my contribution.

Dr. H. Fröhlich (*Bristol*) (*communicated*): In reply to Dr. Böttcher's enquiry, I should like to mention that the condition $\epsilon < 9$ was derived by Sack and myself[6] by showing that Onsager's formula constitutes the zero approximation in a theory in which ϵ is developed in a power series of a parameter γ. Here $\gamma = \mu R/kT$, where R is the reaction field. Using the condition $\gamma \leq 1$ and the value for R given by Onsager leads to $\epsilon \lesssim 9$. But it seems worth mentioning that the zero order approximation (*i.e.* Onsager's formula) holds even if μR is replaced by a different interaction, *e.g.* if it contains terms which are of a non-dipolar nature. In the paper just mentioned we have also calculated the next (*i.e.* first) approximation in γ, and we found (*cf.* § 7) that this leads to multiplication of the zero approximation for ϵ by a factor $(1 + \gamma/2)$ in the case of solids, and by $(1 + \gamma/24)$ in the case of liquids.* This rather suggests that although for solids Onsager's approximation is rather limited, it might be valid for liquids even beyond $\gamma = 1$, *i.e.* even for $\epsilon > 9$ as Dr. Böttcher suggests from the experimental evidence. A value of $\epsilon = 40$ would require $\gamma = 6$, and thus will lead to a small (25 %) first order correction in the case of liquids. I feel, however, that the theoretical foundation for such an extension is not very safe, since nothing is known about higher approximations. For this purpose a deeper knowledge of the theory of liquids than is available at present would be required. It rather seems to me that a way to make further progress should be to make use of Kirkwood's formula for the dielectric constant of liquids to discuss the interaction between dipolar molecules in a semi-empirical way.

---

[3] van Arkel and Snoek, *Physik. Z.*, 1932, **33**, 662; 1934, **35**, 187.

[4] Böttcher, *Rec. trav. chem.*, 1943, **62**, 120, Table I.

[5] Böttcher, *Physica*, 1939, **6**, 59.

[6] *Proc. Roy. Soc. A*, 1944, **182**, 388.

* It seems that our paper was not available to Dr. Böttcher when he wrote his communication.

Dr. F. C. Frank (*London*) (*communicated*) : I am in full agreement with Dr. Böttcher's final paragraph. What I do mean is that the Mosotti formula is the basic simple law for this case, and that deviations from it are specific, and significant, in just the same way as deviations from the ideal gas laws are taken to be significant, giving us specific information about molecular sizes and attractions rather than challenging the derivation of the ideal laws. And the purport of my paper was to suggest that it can still be useful to consider the dielectric behaviour of substances containing polar molecules in terms of deviations from the Mosotti-Debye theory as an ideal. This attitude does not imply that other approaches may not also be valid.

B. THEORY OF RELAXATION TIMES.

EXPERIMENTAL CURVES OF LOSSES AND OF ANOMALOUS DISPERSION AS A BASIS OF A SPECTRAL METHOD.

By Pierre Girard and Paul Abadie.

Received in French 20th February, 1946, and translated by Dr. M. Welch.

The molecular interpretation given by Debye, of the phenomenon of anomalous dispersion discovered by Drude, is still the most valuable guide to the experimenter.

This theory assumes an ideal medium where the polar molecules are all identical, have spherical shape, and are separated from each other.

When one traces the experimental curves of dispersion and of losses, extending over the whole interval of frequencies corresponding to the dispersion or the absorption (or at least over long intervals of frequencies), important divergencies with the theoretical curves are recorded.

These divergencies arise from the complexity of the real facts. Such molecular interpretation, in following the model of orientation polarisation, leads not only to theories of structure, but also gives information on different processes, including molecular interactions, different modes of binding of water, etc. We will briefly analyse some experimental examples.[1]

Crystalloid Molecules.

Pure Liquids.—The heavy lines of Fig. 1 represent the experimental curves of dispersion and of absorption of a normal long-chain alcohol, octyl alcohol (between 2·8 cm. and about 20 metres) ; the fine line represents the theoretical curve.

We will analyse the experimental dispersion curve. Although the theoretical curve shows a single region of dispersion and a single relaxation time, the experimental curve shows two distinct regions of dispersion, separated by a plateau, and two relaxation times.

---

[1] The results reproduced here were obtained during the years 1940-45.

Several interpretations of this divergence, are possible. The existence of associated polar molecules (2 to 2 for example) can be assumed along with simple molecules. But this explanation fails to account for a number of experimental data. For example (see Fig. 2) dilution, which ought to break up the associated molecules, makes a third region of dispersion appear, which cannot reasonably be attributed to a new mode of association (3 to 3 for example).

The most acceptable interpretation, to our mind, brings in the shape of the molecule. The Debye theory, which assumes spheres rotating in the alternating field, implies a single region of dispersion and a single relaxation time (τ). However, the octyl alcohol molecule is elongated. In a memoir which extends the theory of dielectric dispersion to the ellipsoidal model, Francis Perrin [2] establishes that although one single τ only is possible in the theory of spherical molecules, there are three values of τ possible on the theory of an ellipsoid with three axes, but only two values of τ for the ellipsoid of revolution.

The geometrical model suggested by the experimental curve of dispersion of Fig. 1 is an ellipsoid of revolution. The oscillation of such a polar ellipsoidal molecule in an alternating field is complex. It is interesting to analyse it.

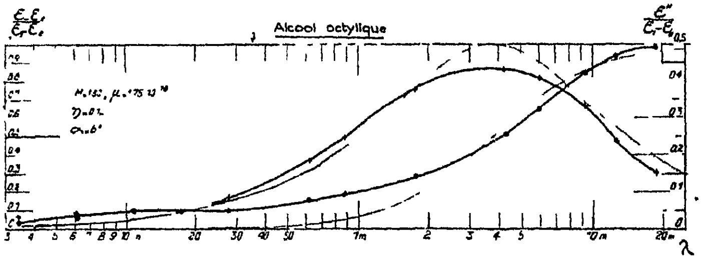


FIG. 1.

Briefly, it can be analysed into simple oscillations, each being determined by a component of a permanent moment about a geometrical axis.

The oscillation involving the least friction corresponds to the first dispersion region and to the smallest value of τ; this is easy to visualise, consisting of a rotation of the ellipsoid about the major axis. Oscillations of the major axis itself in an indeterminate plane, correspond to the second dispersion region and the largest τ.

This brief analysis elucidates the complexity of notions of friction and viscosity when looked at on a molecular scale. If molecular obstructions are taken into account, for pure liquids, the effect of friction is very different, the ellipsoids of revolution being pressed against each other, rotating about or oscillating along the major axis. The effects, although very small in the first case, are considerable in the second. In practice there will be as many viscosity effects as there are modes of oscillation.

Under these conditions it is evident that the term η which in the expression for τ is related to the molecular viscosity

$$\tau = 4\pi\eta a^3/kT$$

cannot be evaluated by a macroscopic measurement of viscosity (rate of outflow from a capillary or velocity of fall of a sphere). Also experience shows that the experimental value always differs, often considerably, from the theoretical value of τ.

---

[2] Perrin, *J. Phys. et Rad.*, 1934, 5, 497.

There is a second disagreement between theory and experiment. In order to facilitate comparison by bringing the calculated curve of Fig. 1 into the same region of abscissa as the experimental curve, we must attribute a value different from the real value to the radius of the molecule.

Let us add that all the normal alcohols give rise to the same type of loss curve, the only difference being the spreading out of the curves, increasing proportionally with the length of the carbon chain.

Crystalloids.

Solutions.—The geometrical form which we attribute to the molecules of normal alcohols is far from being rigid. Fig. 2 shows the extent to which the form of an octyl alcohol molecule depends on its molecular environment.

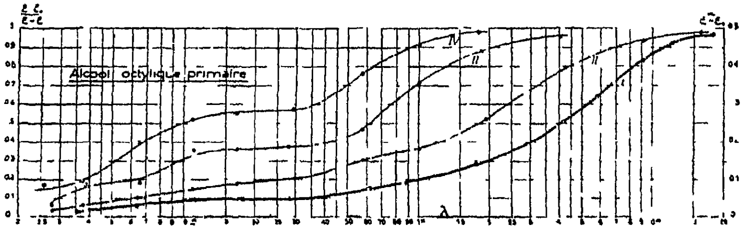


FIG. 2.

In Fig. 2 curve I is the dispersion curve of pure octyl alcohol, curve II that of the same alcohol diluted with an equal part of liquid vaseline. Curve III diluted with an equal part of hexane, and curve IV in benzene.

These various diluents are not polar, and give rise neither to dispersion nor to absorption. The dispersion obtained is therefore due to the alcohol only.

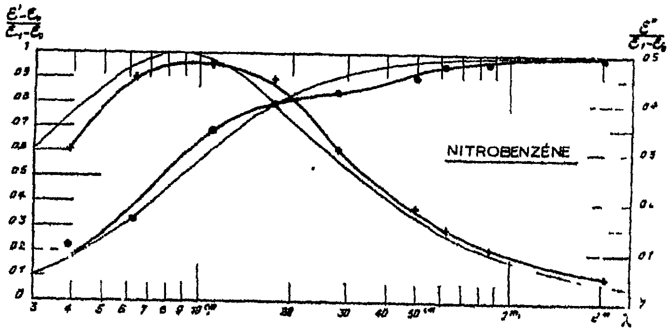


FIG. 3.

Now, as the dispersion curves of the alcohol in these solutions differ from the curve of the pure alcohol, the deformation must vary with the solvent. For these three solutions we see the more or less well defined appearance of 3 dispersion regions and 3 relaxation times, in place of the 2 dispersion regions characteristic of the pure alcohol. The geometrical model changes; it passes from the ellipsoid of revolution to the ellipsoid of 3 axes.

These deformations demonstrate the plasticity of alcohol molecules. However, the strongest interest aroused by such deformations, is the possibility of detecting interactions between molecules of different kinds in the body of the solution. In this, we have a precise method of penetrating directly into the field of molecular interactions, still so little explored.

We have said of the form of the dispersion curve that its outline changes with the geometrical form of the volume circumscribed in space by the molecule. Fig. 3 shows one for a flat disc (nitrobenzene). Two dispersion regions can still be clearly recognised, but this time the first dispersion curve has very much the higher ordinate. Carefully repeated measurements affirm the existence of the second dispersion curve. We will limit ourselves to these examples relative to crystalloid molecules[3] and will now deal with colloidal media.

Colloidal Media.

We have especially utilised for the analysis the loss-angle curves of colloidal media in the form of a solid sheet, film or powder. Many of these curves were drawn up over a very large region of λ. Their outline, instead of being bell-shaped, is much more complicated and irregular.

In fact we have to make not a single curve, but a set of curves, more or less joined and extending over a very large λ interval. Each component curve corresponds to a class of dipoles characterised by their value of τ which progressively increases towards large values of λ. The resulting curve gives us information regarding the distribution of these various categories of dipoles; that is to say, on the dipolar structure of the medium used in the experiment.

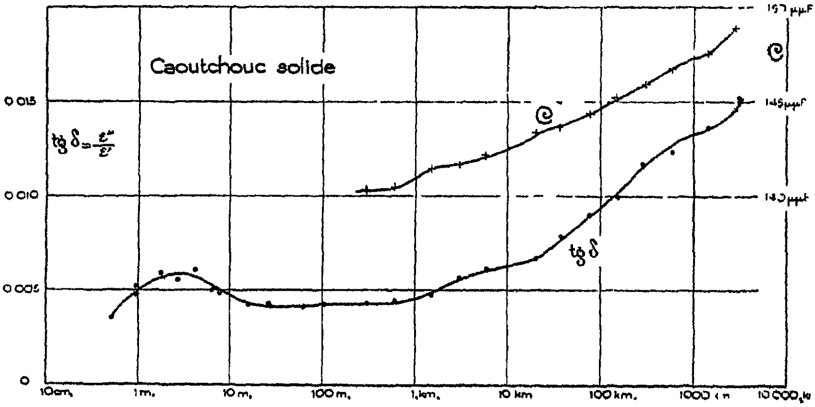


FIG. 4.

A detailed analysis cannot be given by joining up the curves: the position of the segments on the abscissa axis as a relation of the τ value of the corresponding dipoles, and the height of the ordinate as a function of the concentration of the polar groupings and of the magnitude of their μ, will be noted. The outline of the curve will permit the characterisation, in broad lines, of the different categories of polar groupings with regard to their relaxation times.

By comparing the loss curves of a sufficiently large number of colloidal substances, it is seen that each has its own curve and consequently its own characteristic distribution of τ.

Fig. 4 shows the loss curve of a non-vulcanised rubber (without added carbon). Its outline is much more complicated than that of the theoretical Debye curve (bell-shaped), and extends over a very much larger interval of λ. Only a segment (between $\lambda = 0{\cdot}50$ m. and $\lambda = 8$ km.) is shown here.

---

[3] See Girard and Abadie, *Bull. Soc. Chim. de France*, 1945, 12, 207.

In fact we have to make, not a single curve corresponding to identical dipoles (as in the case of polar crystalloidal molecules), but a series of curves, each one corresponding to a category of polar groupings, for which the value of τ increases progressively.

The general outline of the curve is from left to right: a bell-shaped curve corresponding to polar groupings having similar τ values; then a flat and drawn out section corresponding to a succession of categories of dipoles of similar concentration and having values of τ which increase in a continuous manner; then towards large values of λ and higher values of τ there is a series of bell-shaped curves which overlap, each curve corresponding to chains of similar length.

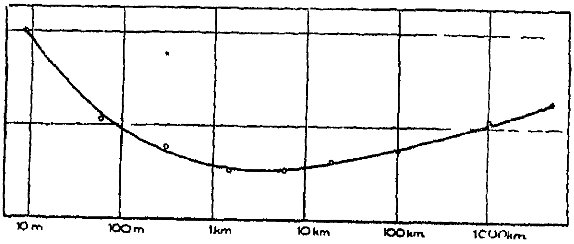


FIG. 5.

A purified rubber, freed from its proteins and resins gives a much simpler curve (Fig. 5).

Fig. 6 refers to five specimens of natural non-vulcanised rubber, with different resistances to rupture.

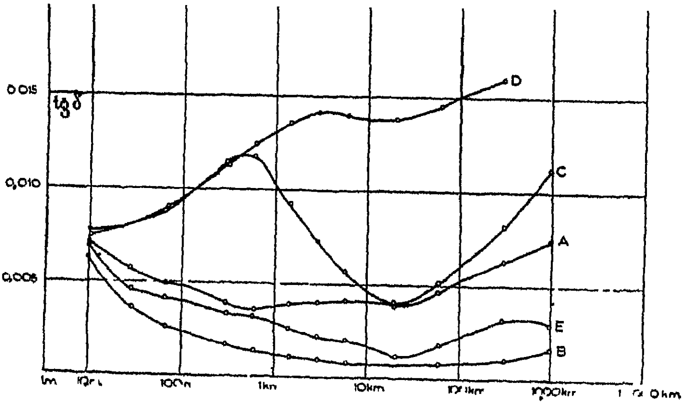


FIG. 6.

They come from the same rubber subjected to different physical treatments (maceration, rolling, heat) which modify their properties. In D the concentration of long chains explains the height of the ordinate towards large λ's and also its very good resistance to rupture. Specimens B and E which have similar resistances to rupture, which, although still good, are inferior to that of D, and are characterised by categories of polar groupings of which the concentrations are sensibly the same for each category and for which the values of τ (and the lengths of the chains) increase in a continuous manner.

Specimen A, although a little different, is closely related both in resistance and structure to B and E.

The curve for C refers to a bad rubber, with a feeble resistance to rupture, in which the values of τ distribute themselves into two sufficiently homogeneous and quite distinct groups, the mean values of τ of each group differing considerably among themselves. This sample is interesting as it indicates the existence of a relation between dipolar structure of bodies and their physical properties.

Figs. 7 and 8 refer to nitro- and aceto-cellulose (films) at different temperatures.

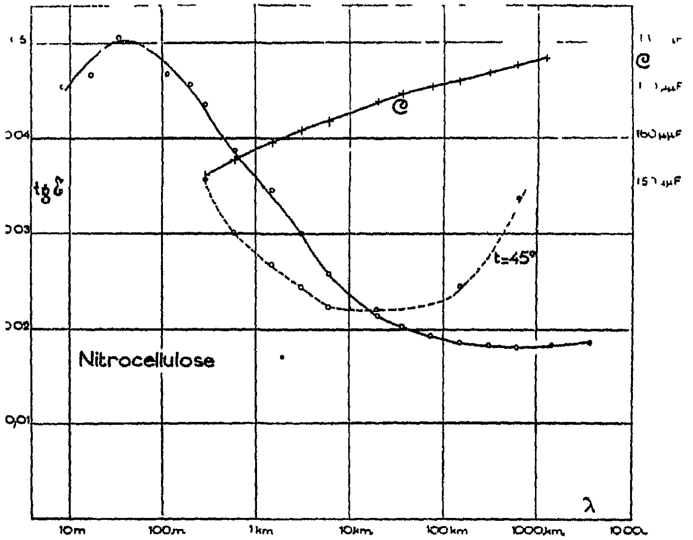


FIG. 7.

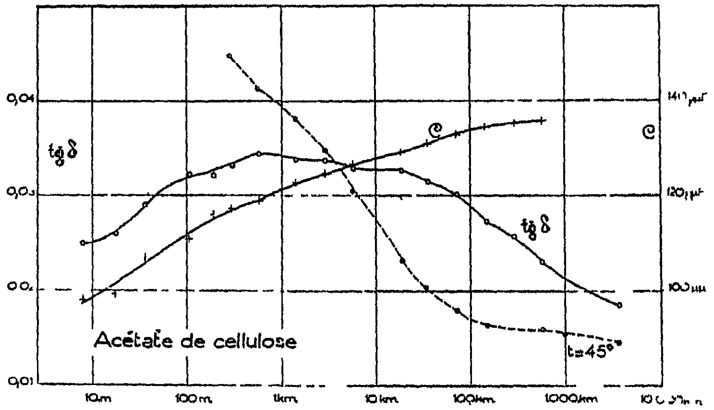


FIG. 8.

The loss curve of aceto-cellulose (small circles joined by an unbroken line) is appreciably different at 18° from that of nitrocellulose at the same temperature. The curve at 45° (small circles joined by a broken line) shows not only a displacement towards short λ's, and consequent diminution of the viscosity effect, but also an important change of form ; the curve approaches that of nitrocellulose at 18°. Because of this sensitivity to variation of viscosity, these films, and similarly a sheet of rubber, behave like a liquid medium, at least in respect to their dielectric properties.

To each temperature there corresponds a characteristic dipolar structure, that is, a different substance. We can thus understand the importance of recording either the losses or the dielectric constants of these substances, at different λ's and over as long an interval of λ as possible, the temperature being fixed. This is the only possible basis for use of a spectral method.

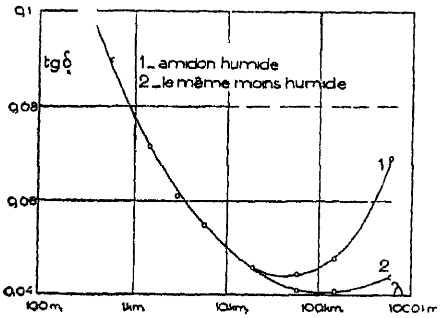


Fig. 9.

Free Water and Bound Water.

The method is applicable to solids, not only in the form of films, leaves or sheets, but also to powders, certain precautions being taken regarding the similarity of the grains, interstitial air, etc. The application of the method to powders allows a new approach to the old question of free and bound water.

Fig. 9 shows a segment of a curve for starch between λ = 300 m. and λ = 1000 Km. Towards λ = 100 Km. for curve no. 1, *i.e.* starch with 10 % water the rapid change (and the rapidity of variation of capacity) indicates that a conductivity curve has detached itself from the curve of dielectric losses. Curve no. 2 corresponds to the same starch partially dehydrated *in vacuo* over P_2O_5. The two conductivity curves join up to the same curve of dielectric losses.

It is interesting to compare with this starch spectrum that of gluten, the second constituent of flour, at the same temperature as above, and likewise in two different hygroscopic states. In Fig. 10 we see, towards large values of λ but at different levels of abscissa, two conductivity curves. In the case of curve I, that is for the wetted gluten (10 % water), the curve of dielectric losses changes towards λ = 10 Km. and its gradient becomes steeper. In the case of curve II, the drier gluten (3 % water), is beyond λ = 100 Km. and its gradient becomes not so steep. However, as we have already seen, whereas the conductivity curves for starch join the same curve of dielectric losses, in this case there are two distinct loss curves, differing greatly in height. As the same powder is concerned, the difference of losses can only be attributed to a difference of polarisability, *i.e.* for the curve of the stronger losses, to an increase of the μ value of the different polar groupings of gluten.

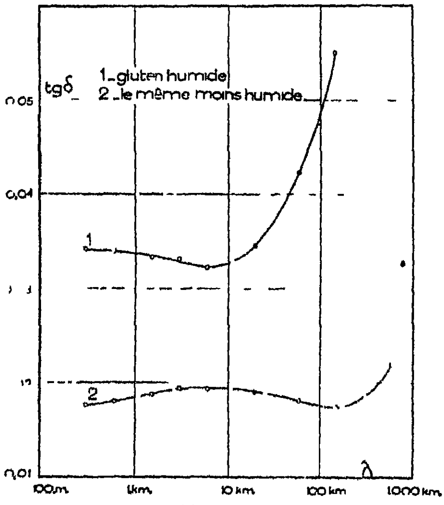


Fig. 10.

Thus, under the experimental conditions, the water is retained in gluten in two distinct forms: free water to which can be attributed the conductivity curves towards long λ's; and combined water, this linkage being characterised by the fact that the water contributes to the increase of the μ of the polar groupings of gluten. It is this chemically active combined water which plays its part in the process of hydrolysis. Under conditions of humidity and temperature which are a little different, starch also retains water in these two forms and not only as free water.

Laboratoire National de Radioelectricité,
196 *rue de Paris,*
Bagneux (Seine),
France.

POWER FACTOR AND TEMPERATURE COEFFICIENT OF SOLID (AMORPHOUS) DIELECTRICS.

BY M. GEVERS AND F. K. DU PRÉ.

Received 8th April, 1946.

1. Introduction.

Most of the technically important solid dielectrics are either of the amorphous type, *e.g.* natural and synthetic resins, glasses, ebonite, mycalex, etc., or they consist of a crystalline mass containing many lattice irregularities and impurities, *e.g.* the group of ceramic dielectrics, especially those prepared with addition of plastifying materials or flux compounds to the substantive material. For sake of convenience both groups will be indicated as " amorphous dielectrics ".

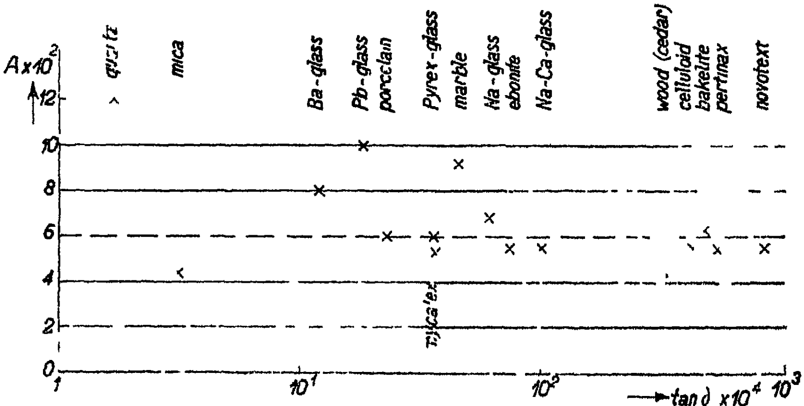


FIG. 1.—The ratio $A = \dfrac{1/\epsilon' \, . \, \dfrac{\partial T}{\partial \epsilon'}}{\tan \delta}$ for various amorphous dielectrics.

They generally have the following properties :

(*a*) the dielectric constant ϵ' varies only slightly with frequency ;
(*b*) the power factor tan δ is also almost independent of the frequency, generally decreasing gradually with increasing frequency.

We have measured the dielectric constant and the power factor of a large number of these amorphous dielectrics in a wide range of temperatures and frequencies * and found a third property shown by all these substances :

(c) the ratio A of the temperature coefficient of the dielectric constant, $(1/\epsilon')\frac{\partial \epsilon'}{\partial T}$, to the power factor tan δ has nearly the same value for all dielectrics having moderate values of dielectric constant and coefficient of thermal expansion. This property is illustrated by Fig. 1, showing the ratio A for different dielectrics. From this it is seen that the magnitude of A equals about 0·06.

In the present paper we show how these properties, and especially the rather remarkable law mentioned under (c), can generally be understood on the basis of the peculiar structural properties of more or less amorphous solids.

2. Basic Picture.

We assume as a starting point that part of the dielectric constant of these materials, and their power factor, are a consequence of the presence of dipoles and (or) ions in the dielectric. The electrical dipoles are either due to the polar groups present in the molecules building up the amorphous substance or to occluded or impurity polar molecules (low polymer molecules, water, etc.), which are both subject to more or less restricted rotation. The ions may be either ions originating from or belonging to the substance itself (*e.g.* in the case of heteropolar inorganic compounds) or ions formed by electrolytic impurities, which can both travel over certain distances through the structure of the amorphous solid, along crystal flaws or irregularities, etc. In the case of a glass, for instance, we must assume that apart from the ions which are responsible for the electrolytic conductivity many ions can move to and fro over a restricted region formed by the interstices in the SiO_2 skeleton.

A dipole will be bound by the molecular field of the surroundings to certain positions and can only alter its position if a certain activation energy q is supplied. In the same way a displacement of ions involves an activation energy because the motion of these ions is counteracted by the surrounding ions or atoms.

Now it is in accordance with the nature of an amorphous or irregularly formed solid that these activation energies will differ for different ions or dipoles. Also for a particular ion the field of force in which the ion travels will be rather irregular and accordingly the potential barriers will vary from one point to another; analogous conditions may be expected for a rotating dipole.

Accordingly we base our theory on the assumption of a statistical scattering of the activation energies determining the progressive or rotational diffusion of ions or dipoles.† Introducing very simple assumptions as to the distribution law of these activation energies, we can then easily explain the characteristic properties of amorphous dielectrics, and we find quantitative agreement in the case of property (c), as the theory leads to a value of A of the right order of magnitude.

3. Dielectric Constant and Power Factor.

We first consider briefly the case that all particles (ions or dipoles) have the same activation energy. In that case the theory follows lines already given by the well known considerations of Wagner and Debye; in section 4 the theory will be generalised to the case of a distribution of activation energies.

---

* The detailed theory and results of experiments will be published in *Philips Research Reports*.

† In order to explain the properties (a) and (b) von Schweidler and Wagner have already introduced the assumption of a distribution of relaxation times. This assumption is physically less reasonable than that made above and it is actually not possible to explain the third property with the aid of this theory without introducing additional special suppositions.

When a constant electric field is applied to the dielectric it is polarised almost immediately as a consequence of the displacement of the electrons in the atoms. This polarisation at the time t_0 may be called p_∞ (polarisation for very large frequencies). In addition we have now gradually increasing polarisation due to the orientation of the dipoles or the displacement of the ions, which approaches exponentially the final value $p_0 - p_\infty$ (see Fig. 2), so that at any moment the additioned polarisation will be

$$p_a = (p_0 - p_\infty)\left(1 - e^{-\frac{t-t_0}{\tau}}\right) \quad . \quad . \quad . \quad . \quad (1)$$

where τ is a time constant or relaxation time. This polarisation is completely analogous to a diffusion process and the time constant τ, determining the rate of polarisation, will be given by

$$\tau = \tau_0 e^{q/kT} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where q is the activation energy causing the frictional force opposing the displacement or rotation of the particles. This energy must be supplied by thermal motion and the proportionality factor τ_0 will be of the order of magnitude of the vibration time of the particles in their potential box, hence $\approx 10^{-13}$ sec.

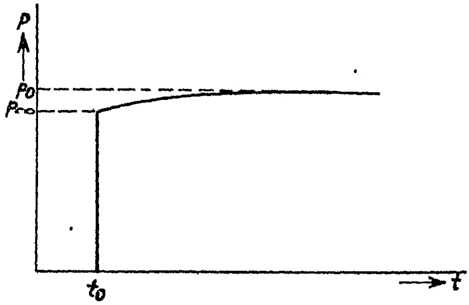


FIG. 2.—Polarisation p as a function of the time t when at the instant t_0 a constant electric field is applied to the dielectric.

When an alternating electric field is applied to the dielectric, the polarisation will depend on the frequency ω and be out of phase with the electric field; this phase difference δ is the loss angle of the dielectric, the quantity $\tan \delta$ the power factor. This phase lag is due to the slow polarisation mechanism discussed above and gives rise to a heat dissipation in the dielectric, which is proportional to $\tan \delta$ and the frequency ω.

In symbolic calculus, an electric field of sine wave may be written as $E = E_0 e^{j\omega t}$, where $j = \sqrt{-1}$. It may be shown, by application of Hopkinson's principle of superposition, that the polarisation p as a function of the frequency is given by the complex function

$$p = p_\infty + \frac{p_0 - p_\infty}{1 + j\omega\tau} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

at the same time giving the real part of p and the phase difference with the electric field (from the imaginary part). This equation may be written

$$p = p_\infty + \frac{p_0 - p_\infty}{1 + \omega^2\tau^2} - j(p_0 - p_\infty)\frac{\omega\tau}{1 + \omega^2\tau^2} \quad . \quad . \quad (4)$$

because of the relation $p = \frac{\epsilon - 1}{4\pi}E$, we may also write

$$\epsilon = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + \omega^2\tau^2} - j(\epsilon_0 - \epsilon_\infty)\frac{\omega\tau}{1 + \omega^2\tau^2} \quad . \quad . \quad (5)$$

Denoting the real part by ϵ' and the imaginary part by ϵ'', from the above equation we obtain:

$$\epsilon' = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + \omega^2\tau^2} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

being the measured value of the dielectric constant, and

$$\epsilon''/\epsilon' = \tan\delta \approx \frac{\epsilon_0 - \epsilon_\infty}{\epsilon_\infty} \frac{\omega\tau}{1 + \omega^2\tau^2} {}^* \quad . \quad . \quad . \quad (7)$$

assuming that the difference between ϵ_0 and ϵ_∞ is small, as will generally be the case.

From the above equations and as shown by Fig. 3, the dielectric constant decreases with increasing frequency, whereas the power factor has a maximum value at the frequency $\omega = 1/\tau$. At low frequencies the ions and dipoles can contribute to the full extent to the dielectric constant without time lag. At very high frequencies the motion or orientation of the particles cannot follow the alternating field and accordingly does not contribute to the polarisation. At the frequency corresponding with the relaxation time the phase difference has a maximum value.

The dependence of ϵ' and $\tan\delta$ on the temperature may be found by substituting $\tau = \tau_0 e^{q/kT}$ (2). The contribution $\epsilon_0 - \epsilon_\infty$ of the ions and dipoles to the dielectric constant, may be expressed by their susceptibility χ, which may be defined by

$$- \epsilon_\infty = 4\pi n\chi \quad . \quad (8)$$

where n is the number of ions and dipoles in unit of volume. Eqn. (8) is analogous to the relation between the dielectric constant ϵ' of a diluted gas and the susceptibility of its molecules:

$$\epsilon' - 1 = 4\pi n\chi.$$

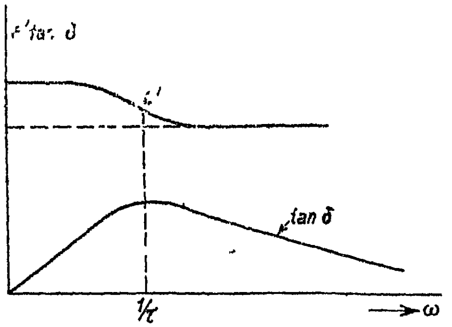


FIG. 3.—Dielectric constant ϵ' and power factor $\tan\delta$ as a function of the frequency ω.

4. Dielectric Constant, Power Factor and Temperature Coefficient in the case of a Distribution of Activation Energies.

Let the distribution function be denoted by $G(q)$, which may be a continuous function in relation with the large number of ions and dipoles, for instance the function illustrated by Fig. 4. When n is the total number of particles per cm.^{3} contributing to the dielectric losses, the number of particles having activation energies between q and $q + dq$ amounts to $nG(q)dq$. It will be clear that the total area of the curve $G(q)$ equals unity, because of the condition, that the integration over all values of q gives the total number per cm.^{3}, so

$$n \int_0^\infty G(q)\,dq = n. \quad . \quad . \quad . \quad . \quad . \quad (9)$$

For our purpose it is not essential to give an explicit function for $G(q)$; only some assumptions must be made with respect to the shape of the curve. It is in accordance with the properties of amorphous dielectric solids to assume that the curve is rather flat, so $G(q)$ increases slightly for small values of q and decreases slightly at large values of q.

---

* See, *e.g.* Schupp, *Wiss. Veröff. Siemens*, 1938, 17, 1.

In deriving the expressions for the dielectric constant and power factor in the case of the distribution law $G(q)$ of the activation energies from the equations given in the preceding section, we allow for the fact that the polarisation of a particle may slightly depend on the position of the ion or dipole in the field of forces and therefore on the activation energy q. Generally instead of the susceptibility χ we will therefore write $\chi(q)$, indicating that the susceptibility may be a function of q. We assume that $\chi(q)$ does not depend on T explicitly, but on q only. With this assumption we are in accordance with Debye's theory on high viscous solids.[1]

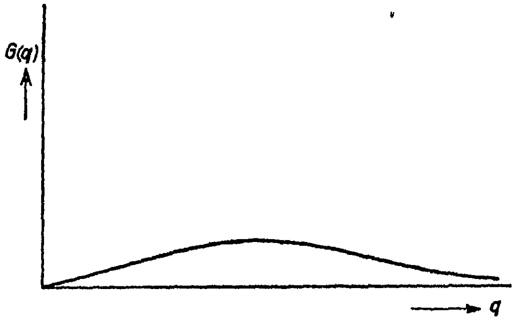


FIG. 4.—Distribution function of the activation energies q.

The equation for the dielectric constant and power factor may be obtained from the simple expressions (6) and (7) where $\epsilon_0 - \epsilon_\infty$ is given by eqn. (8)), when n is replaced by $nG(q)\mathrm{d}q$, χ by $\chi(q)$ and by integrating over all values of q. Accordingly from (6) and (8) we obtain for the dielectric constant

$$\epsilon' = \epsilon_\infty + 4\pi n \int_{q=0}^{\infty} \frac{G(q)\chi(q)}{1+\omega^2\tau^2}\mathrm{d}q \quad . \quad . \quad . \quad (10)$$

and for the power factor from (7) and (8),

$$\tan\delta = \frac{4\pi n}{\epsilon_\infty}\int_0^\infty G(q)\chi(q)\frac{\omega\tau}{1+\omega^2\tau^2}\mathrm{d}q \quad . \quad . \quad . \quad (11)$$

where the relaxation time τ is a function of q, as defined by (2).

The integrals occurring in eqn. (10) and (11) may be simplified strongly owing to the properties of the functions $\frac{1}{1+\omega^2\tau^2}$ and $\frac{\omega\tau}{1+\omega^2\tau^2}$. Substitution of

$$\tau = \tau_0 e^{q/kT}$$

in the former function gives

$$\frac{1}{1+\omega^2\tau^2} = \frac{1}{1+\omega^2\tau^2_0 e^{2q/kT}} \quad (12)$$

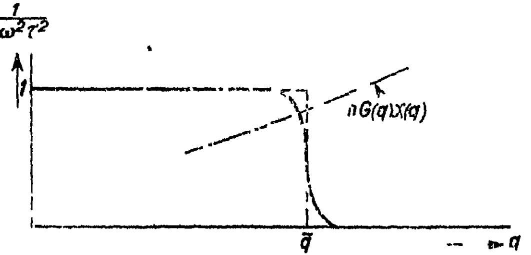


FIG. 5.
$\frac{1}{1+\omega^2\tau^2}$ as a function of the activation energy q.

Now we introduce a certain activation energy $\bar{q}$, being the value of q corresponding with the case $\omega\tau = 1$, so

$$\omega\tau_0 e^{\bar{q}/kT} = 1 \quad \text{or} \quad \omega\tau_0 = e^{-\bar{q}/kT} \quad . \quad . \quad . \quad (13)$$

Substitution of this expression in (12) gives

$$\frac{1}{1+\omega^2\tau^2} = \frac{1}{1+e^{2(q-\bar{q})/kT}} \quad . \quad . \quad . \quad (14)$$

---

[1] Debye, *Physik. Z.*, 1935, 36, 100, 193; *Bull. Ac. Roy. de B.*, 1935, 2, 166; Debye-Ramm, *Ann. Physik*, 1937, 28, 28.

this function is shown by the full curve in Fig. 5; because of the sharp decrease at $q = q$, in the neighbourhood of $\bar{q}$[2] the curve may be approximated by the dotted line.

So for $q < \bar{q}$ the function under consideration equals 1, whereas for $q > \bar{q}$ the function equals zero. In order to evaluate the integral occurring in (10) this function must be multiplied by $nG(q)\chi(q)$. When the variation of this function remains small with respect to the variation of $\frac{1}{1 + \omega^2\tau^2}$ in the neighbourhood of $q = \bar{q}$, then the following approximation may be applied:

for $q \leq \bar{q}$ $$\frac{1}{1 + \omega^2\tau^2}G(q)\chi(q) = G(q)\chi(q)$$

whereas for $q > \bar{q}$ $$\frac{1}{1 + \omega^2\tau^2}G(q)\chi(q) = 0.$$

From this it will be clear that the following approximation is permitted

$$\int_0^\infty \frac{G(q)\chi(q)}{1 + \omega^2\tau^2}\,dq = \int_0^{\bar{q}} G(q)\chi(q)\,dq$$

and therefore the dielectric constant ϵ' is given by

$$\epsilon' = \epsilon_\infty + 4\pi n \int_0^{\bar{q}} G(q)\chi(q)\,dq. \quad . \quad . \quad . \quad (15)$$

Substitution of $\tau = \tau_0\,e^{q/kT}$ in $\frac{\omega\tau}{1 + \omega^2\tau^2}$ gives

$$\frac{\omega\tau}{1 + \omega^2\tau^2} = \frac{\omega\tau_0\,e^{q/kT}}{1 + \omega^2\tau_0^2\,e^{2q/kT}}$$

or, introducing again the characteristic activation energy $\bar{q}$, as given by (13),

$$\frac{\omega\tau}{1 + \omega^2\tau^2} = \frac{e^{(q - \bar{q})/kT}}{1 + e^{2(q - \bar{q})/kT}}$$

This function, as shown by Fig. 6, has a sharp maximum at $q = \bar{q}$.

Because of the slight variation of $G(q)\chi(q)$ and the rapid variation of $\frac{\omega\tau}{1 + \omega^2\tau^2}$, the function $G(q)\chi(q)$ may be supposed to have the constant value $G(\bar{q})\chi(\bar{q})$ in the region where $\frac{\omega\tau}{1 + \omega^2\tau^2}$ is unequal to zero and therefore $\int_0^\infty G(q)\chi(q)\frac{\omega\tau}{1 + \omega^2\tau^2}\,dq$ may be approximated by

$$G(\bar{q})\chi(\bar{q}) \int_0^\infty \frac{\omega\tau}{1 + \omega^2\tau^2}\,dq.$$

Evaluation of $$\int_0^\infty \frac{\omega\tau}{1 + \omega^2\tau^2}\,dq.$$

(see Appendix I) gives $\pi/2kT$ and therefore

$$\tan\delta \approx \frac{2\pi^2}{\epsilon_\infty}nG(\bar{q})\chi(\bar{q})kT \quad . \quad . \quad . \quad . \quad (16)$$

Physically the expressions for ϵ' (15) and $\tan\delta$ (16) may be interpreted as follows. To the *dielectric constant* of the amorphous dielectric as a first approximation only those ions and dipoles contribute for which the activation energies at the actual temperature T do not exceed the value $\bar{q}$, *i.e.* whose relaxation times exceed the value $\tau = 1/\omega$, where ω is the angular frequency of the alternating electric field. As for the *dielectric losses* the same result is obtained as in the case that only those particles contribute to the losses for which the activation energies are equal to $\bar{q}$, *i.e.* whose relaxation times at the temperature T of the dielectric equal $1/\omega$.

---

[2] $\bar{q}$ is of the order of magnitude 10 to 20 kT.

From the equations obtained, the dielectric properties mentioned in the first section may be explained as follows :—

(*a*) The dielectric constant ϵ' varies only slightly with frequency. Eqn. (15) shows that the part of ϵ' which is dependent on the frequency, is a function of the quantity $\bar{q}$ only. From (13) it follows that

$$\bar{q} = kT \ln 1/\omega\tau_0 \quad (17)$$

which varies only slightly with variation of the frequency. This is shown by Table I, indicating the value of

$$\bar{q}/kT = \ln 1/\omega\tau_0$$

at different frequencies for the case

$$\tau_0 = 10^{-14} \text{ sec.}$$

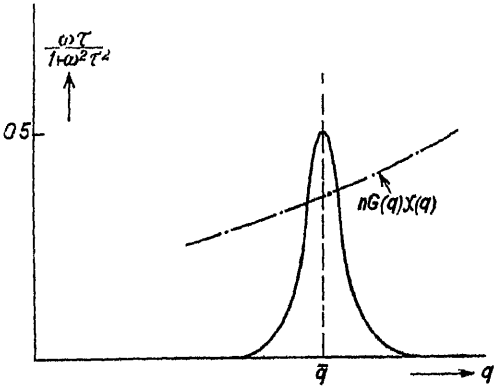


FIG. 6.

$\frac{\omega\tau}{1 + \omega^2\tau^2}$ as a function of the activation energy q.

Since the value of the integral $\int_0^{\bar{q}} G(q)\chi(q)\mathrm{d}q$ depends only on $\bar{q}$, which varies slightly with ω, it may be seen from Table I that ϵ' decreases only slightly with increasing frequency. The dependence of ϵ' on frequency may be expressed by the following relation

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial\omega} = \frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial\bar{q}}\frac{\partial\bar{q}}{\partial\omega} = \frac{-4\pi}{\epsilon'} nG(\bar{q})\chi(\bar{q})\frac{kT^*}{\omega} = -\frac{2}{\pi}\frac{\tan\,\delta}{\omega}$$

hence

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial \ln \omega} = -\frac{2}{\pi}\tan\,\delta \quad . \quad . \quad . \quad . \quad (18)$$

In the case of a dielectric, having $\tan\,\delta = 10 \times 10^{-4}$, the relative variation of the dielectric constant between the angular frequencies 10^3 and 10^9 ($\Delta \ln \omega = 2{\cdot}3 \times 6 = 13{\cdot}8$) amounts to

$$\Delta\epsilon'/\epsilon' = -13{\cdot}8 \times 2/\pi \times 10 \times 10^{-4} = -90 \times 10^{-4}.$$

or less than 1 % in the large frequency range mentioned.

TABLE I.—$\tau = 10^{-14}$ sec.

ω (Rad. Sec.$^{-1}$).	$\omega\tau_0$ (Rad.).	$\bar{q}/kT = \ln 1/\omega\tau$.
10^3	10^{-11}	25·4
10^5	10^{-9}	20·7
10^7	10^{-7}	16·1
10^9	10^{-5}	11·5

(*b*) The power factor $\tan\,\delta$ is almost independent of the frequency. From eqn. (16) it follows that $\tan\,\delta$ is proportional to $G(\bar{q})\chi(\bar{q})$, according to the assumption made being a slightly variable function of the characteristic activation energy $\bar{q}$, and therefore $\tan\,\delta$ will be a slight function of ....

(*c*) The temperature coefficient of the dielectric constant, defined by $1/\epsilon' . \partial\epsilon'/\partial T$ may be derived from expression (15) for ϵ' by differentiation with respect to the temperature, occurring implicitly in the quantity $\bar{q}$, as given by (17).

---

* The value of $\partial\epsilon'/\partial\bar{q}$ may be obtained by differentiation of the integral, occurring in eqn. (15) to the upper limit ; the result is the value of the integrand in which the variable is equal to the upper limit $\bar{q}$.

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial T} = \frac{1}{\epsilon'}\frac{\partial\epsilon_\infty}{\partial T} + 4\pi n\frac{\partial}{\partial T}\int_0^{\bar{q}} G(q)\chi(q)\,\mathrm{d}q = \frac{1}{\epsilon'}\frac{\partial\epsilon_\infty}{\partial T} + \frac{4\pi n}{\epsilon'}G(\bar{q})\chi(\bar{q})k\ln 1/\omega\tau_0. \quad (19)$$

To a first approximation the first part $\frac{1}{\epsilon'}\frac{\partial\epsilon_\infty}{\partial T}$ of the temperature coefficient may be obtained from Clausius-Mosotti's law relating the polarisability α of the atoms to the dielectric constant ϵ_∞ :

$$\frac{\epsilon_\infty - 1}{\epsilon_\infty + 2} = \frac{4\pi}{3}n\alpha$$

where α is independent of the temperature.

From this equation is obtained (see Appendix II)

$$\frac{1}{\epsilon'}\frac{\partial\epsilon_\infty}{\partial T} = -\frac{\alpha_l}{\epsilon'}(\epsilon_\infty - 1)(\epsilon_\infty + 2) \approx -\alpha_l\left(1 + \epsilon_\infty - \frac{2}{\epsilon_\infty}\right)$$

where α_l = coefficient of linear thermal expansion of the substance considered. Consequently ϵ_∞ gives rise to a negative part of the temperature coefficient of the dielectric constant. Using eqn. (16) the second part of the temperature coefficient may be written in the form

$$\frac{4\pi n}{\epsilon'}G(\bar{q})\chi(\bar{q})\,k\ln 1/\omega\tau_0 \approx \frac{2}{\pi T}\ln 1/\omega\tau_0\tan\delta.$$

For the total temperature coefficient is found therefore

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial T} = A\tan\delta - \alpha_l\left(1 + \epsilon_\infty - \frac{2}{\epsilon_\infty}\right) \quad . \quad . \quad . \quad (20)$$

where the quantity $$A = \frac{2}{\pi T}\ln 1/\omega\tau_0 \quad . \quad . \quad . \quad . \quad . \quad (21)$$

Except in the case of dielectrics having very low losses and (or) a large value of dielectric constant ϵ_∞ (*e.g.* rutile) or a high thermal expansion (*e.g.* polystyrene), the negative part in (20) may be neglected ; consequently for most amorphous dielectrics is obtained

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial T} = A\tan\delta \quad . \quad . \quad . \quad . \quad . \quad (22)$$

Now it appears that the quantity A, as given by (21) depends only slightly on the frequency and on τ_0, the latter being the only constant of the substance occurring in (22). This may be seen from Table II, indicating the value of A at different frequencies and two widely different values of τ_0 at 300° K. The experimental values of A suggest that usually τ_0 seems to be smaller than 10^{-13} sec. The same has been found in diffusion phenomena. This may be explained by assuming that the activation energy depends on the temperature ; considering only a linear term

$$q = q_0 - aT$$

we have $$\tau = \tau_0 e^{q/kT} = \tau_0\,e^{q/kT - a/k} = \tau'_0\,e^{q/kT}$$

where $$\tau'_0 = \tau_0\,e^{-a/k}.$$

When for instance $aT = 13kT$, as may be in the case of dielectrics having a low melting-point (q_0 being of the order 10 to 20 kT), $\tau'_0 = 10^{-6}\,\tau_0$; in the case of $\tau_0 = 10^{-13}$ sec. $\tau'_0 = 10^{-19}$ sec. Therefore the values $\tau_0 = 10^{-13}$ and 10^{-20} have been taken (τ'_0 again denoted by τ_0) in Table II.

At commercial radio frequencies the value of A runs from about 0·04 to 0·06, nearly independent of the frequency and of the type of the dielectric.

It may be concluded therefore that in the case of amorphous dielectrics, not having very small dielectric losses or extremely large values of the thermal expansion coefficient or of the dielectric constant, *the ratio A between the temperature coefficient of the dielectric constant and the power factor tan δ is approximately independment of the nature of the amorphous dielectric*, being the property (*c*) mentioned in section 1.

For other consequences derived from the theory dealt with in this paper, we refer to the articles which will be published shortly by one of the authors in *Philips Research Reports*.

TABLE II.—$T = 300°$ K.

ω (Rad. Sec.$^{-1}$).	A (Degree^{-1}).	
	$\tau_0 = 10^{-18}$ Sec.	$\tau_0 = 10^{-20}$ Sec.
10^3	0·049	0·083
10^5	0·039	0·073
10^7	0·029	0·064
10^9	0·020	0·054

Appendix

(I)—Evaluation of the integral $\int_0^{q=\infty} \frac{\omega\tau}{1 + \omega^2\tau^2}\,dq$.

From $\tau = \tau_0 e^{q/kT}$ we obtain $q = kT \ln \tau/\tau_0$, hence

$$dq = \frac{kT}{\tau}\,d\tau,$$

therefore
$$\int_0^{q=\infty} \frac{\omega\tau}{1 + \omega^2\tau^2}\,dq = kT \int_{\tau=\tau_0}^{\infty} \frac{\omega d\tau}{1 + \omega^2\tau^2} = kT\,(\tan^{-1}\omega\tau)_{\omega\tau_0}^{\infty}$$
$$= kT\left(\frac{\pi}{2} - \tan^{-1} e^{-\bar{q}/kT}\right)$$

since $\omega\tau_0 = e^{-\bar{q}/kT}$, according to (13).

Now $\tan^{-1} e^{-\bar{q}/kT}$ is very small with respect to $\pi/2$, consequently

$$\int_0^{\infty} \frac{\omega\tau}{1 + \omega^2\tau^2}\,dq \approx \frac{\pi}{2}\,kT$$

to a high approximation.

(II)—Evaluation of $\partial\epsilon_\infty/\partial T$.

From Clausius-Mosotti's law $\frac{\epsilon_\infty - 1}{\epsilon + 2} = \frac{4\pi}{3} n\alpha$, where n is the number of atoms in unit of volume, we obtain by differentiation of both sides with respect to the temperature T:

$$\frac{\partial}{\partial T}\frac{\epsilon_\infty - 1}{\epsilon_\infty + 2} = \frac{4\pi}{3}\alpha n \frac{1}{n}\frac{\partial n}{\partial T} = \frac{\epsilon_\infty - 1}{\epsilon_\infty + 2}\frac{1}{n}\frac{\partial n}{\partial T}$$

or
$$\frac{3}{(\epsilon_\infty + 2)^2}\frac{\partial \epsilon_\infty}{\partial T} = -3\alpha_l \frac{\epsilon_\infty - 1}{\epsilon_\infty + 2}$$

(where α_l = coefficient of linear thermal expansion), hence

$$\frac{\partial \epsilon_\infty}{\partial T} = -\alpha_l(\epsilon_\infty - 1)(\epsilon_\infty + 2).$$

Natuurkundig Laboratorium der N. V. Philipps Gloeilampenfabrieken,
Eindhoven, Holland.

THE DISTRIBUTION OF RELAXATION TIMES IN DIELECTRICS.

By C. G. Garton.

Received 26th March, 1946.

The variation of loss angle with frequency in many dielectrics has not been satisfactorily explained, especially as regards materials having a loss angle nearly constant with respect to frequency. The author presents later in the paper a reasonable hypothesis which accounts for this characteristic. Before discussing this, some preliminary remarks may be helpful.

It is well known that any loss angle, which is independent of field strength and arises from a single time constant, must pass through a maximum when $\omega = 1/\tau$ and fall off as $1/\omega$ and ω respectively at frequencies far removed from this value. If a distribution of time constants exists, covering any limited range, the maximum occurs somewhere within the range, and the loss angle falls in the same manner for any periodic time lying far outside the range of time constants present. Thus, any limited distribution of time constants leads to results not essentially different from the familiar Debye equations. As will be seen later, a constant loss angle can be explained only by a distribution which extends over (and beyond) the whole range of constancy.

These conditions will not be established here. They may be readily deduced from the behaviour of the functions defining the complex dielectric constant, namely for a single time constant,

$$e' \propto 1/(1 + \omega^2\tau^2) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$e'' \propto \omega\tau/(1 + \omega^2\tau^2) \quad . \quad . \quad . \quad . \quad (2)$$

or, for a distribution of time constants,

$$e' \propto \int_a^b \frac{f(\tau)d\tau}{1 + \omega^2\tau^2} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

$$e'' \propto \int_a^b \frac{\omega\tau f(\tau)d\tau}{1 + \omega^2\tau^2} \quad . \quad . \quad . \quad . \quad . \quad (4)$$

It is also well known, but may be worth emphasis, that the variation of loss angle depends only on these functions, and not at all upon the physical mechanism from which the time constant or its distribution arises. Progress can therefore be made only in so far as new models introduce a new distribution function $f(\tau)$. Some modification of the relations given is required if the mechanism, in addition to a relaxation time, possesses a natural frequency of oscillation. The functions involved then become those recently developed (for the optical case) by Van Vleck and Weisskopf [1] and independently (for the dielectric case) by Fröhlich.[2] However, the modification is not material. Although these functions give a slower variation near the maximum than the Debye relation, the variation at periodic times much longer than the natural period involved is still proportional to frequency, as in the Debye case.

We therefore expect that in any frequency range far from a maximum of loss angle, the latter will be found either to rise as ω, or to fall as $1/\omega$.

---

[1] Van Vleck and Weisskopf, *Rev. Mod. Phys.*, 1945, **17**, 97.

[2] Fröhlich, *E.R.A. Report*, Ref. L/T163; 1946, **157**, 478.

The behaviour of many real dielectrics is completely different. It is common to find a loss angle constant (within an order of magnitude) over a frequency range of 10^6 to one, and rising at *both* ends of the range approximately as log ω and $1/\log \omega$ respectively. This behaviour is found in materials as dissimilar as polyethylene, ebonite and mica, and in fact appears characteristic of "good" dielectrics, suggesting that it is due to some mechanism which is always present but is obscured by more important causes of loss, such as the Debye mechanism, in "poor" dielectrics. It cannot be explained, as has been suggested, on the basis of two groups of time constants lying respectively above and below the limits of measurement, since the variation in such a region should have the quite different form $A\omega + B/\omega$.

The existence of a distribution of relaxation times in dielectrics is now generally accepted, and for a more or less broadened maximum of "Debye" type the necessary distributions have been determined [3] by several workers, either (as Yager [4]) in the form of a Gaussian distribution with a disposable dispersion constant, or (as Garton [5]) from the variation of the dielectric constant. This work, however, merely confirms the existence of a distribution, and throws no light on the nature of the mechanism involved.

To achieve the latter, the distribution must be determined from some independent hypothesis concerning the structure of the dielectric, and this must then lead to the observed variation of loss angle. This paper suggests one such hypothesis. The result will be the more convincing if it applies to characteristics differing radically from the "Debye" type, since small deviations from the latter can always be explained as due to random variations of structure in the material. Materials with an approximately constant loss angle are therefore chosen for consideration in the remainder of this paper.

It will first be convenient to determine the types of distribution which result in a constant value of e''. Hereafter, it is always assumed that (as actually observed when the loss angle is constant) the contribution of the energy losing mechanism to the real part of the dielectric constant is negligible, so that the variations of e'' and of tan δ are similar, and we need not further distinguish between them.

Considering equation (4) as a function of frequency, it will be seen by inserting a few trial functions to represent $f(\tau)$, that the only form which can lead to a constant value of e'' over any considerable frequency range is

$$f(\tau) = 1/\tau \quad . \quad . \quad . \quad . \quad . \quad (5)$$

when the solution becomes $$e'' \propto \Big[\arctan \omega\tau\Big]_a^b \quad . \quad . \quad . \quad . \quad (6)$$

which, in the range $\omega a \ll 1 \ll \omega b$ gives substantially

$$e'' \propto \pi/2 = \text{constant}.$$

Assuming now that whatever result we finally obtain will be expressible as a rate process, there will be associated with the problem a function, the form of which we leave open, representing the contribution to the polarisability from any state having an activation energy (depth of potential well) v. This function we write $\delta P \propto F(v)\,\delta v$.

Expressing (4) in the equivalent form

$$e'' \propto \int_a^b \frac{\omega\tau}{1+\omega^2\tau^2}\,F(v)\,\frac{dv}{d\tau}\,d\tau \quad . \quad . \quad . \quad . \quad (7)$$

we observe that to obtain a value of e'' constant with respect to frequency we must have, in accordance with (5)

$$F(v)dv = d\tau/\tau \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

---

[3] Fuoss and Kirkwood, *J. Am. Chem. Soc.*, 1941, 63, 385.

[4] Yager, *Physica*, 1936, 7, 434.

[5] Garton, *J.I.E.E.*, 1939, 85, 625.

which implies (constants being everywhere omitted for brevity)

$$\log \tau = \int F(v)dv \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Equation (9) is a convenient means of ascertaining what types of relations can be associated with a constant loss angle, and certain forms can be rejected immediately by the improbable relations to which they lead. For example, an exponential form for $F(v)$ leads to the improbable consequence ($\log \tau = e^v$), while a Gaussian distribution implies ($\log \tau = \text{erf } v$).

A simple (and as will later be seen, very significant) solution of (8) is

$$F(v) = \text{constant.} \qquad \log \tau = v \quad . \quad . \quad . \quad (10)$$

A distribution in which the population is independent of activation energy appears improbable until it is remembered that $F(v)$ is actually a distribution of polarisability and not of physical elements. (The word "element" is used here and throughout, as a noncommittal alternative for ion or dipole.) It can be expressed as a product

F (population × polarisability per element)

and we observe that an alternative form of (10) is in fact

$$F(v) = F(\tau e^{-v}) = \text{const.} \qquad \log \tau = v \quad . \quad . \quad (11)$$

We have therefore a solution of the problem if an exponential population is associated with a polarisability varying as τ while $\log \tau$ varies as v. Our model will possess these characteristics.

Two results will be required which may easily be deduced as rate processes, though space will not permit the proof to be given. These are:

(*a*) If the time constant of "a well" be defined as that of a system of two equal wells, then the time constant of a system of two unequal wells with individual time constants τ and τ_0 is

$$2\tau\tau_0/(\tau + \tau_0). \quad . \quad . \quad . \quad . \quad (12)$$

(*b*) If an element of constant moment (dipole moving through a given angle, or ion through a given distance) oscillates between two potential wells of unequal depth, its contribution to the polarisability varies as

$$\tau\tau_0/kT(\tau + \tau_0)^2 \quad . \quad . \quad . \quad . \quad (13)$$

The dependence upon time constants in (13) is in the form of a dimensionless ratio, and does not infer that the polarisability varies with time. Equations (12) and (13) could, of course, equally well be expressed in terms of the well depths by making the appropriate substitutions connecting τ and v. It will be observed that relation (13) involves the connection between τ and polarisability which we were led to expect from discussion of equation (11).

We may now introduce a model. Consider first a wholly amorphous substance—in the extreme case, a liquid. There will be no permanent potential wells present other than those involved in the structure of individual molecules. But, owing to the diffusion of molecules past one another (fast in the case of a liquid, but becoming indefinitely slower with increase of viscosity toward a vitreous structure) there will be a population of temporary wells (semi-permanent if the viscosity be sufficient) the depths of which will be randomly distributed according to all possible molecular juxtapositions. Our reason for interest in these wells will become apparent later. In a micro-crystalline material such as polyethylene, the temporary wells will be absent within the crystallites, but will exist at intercrystalline boundaries and in the amorphous material between them. It is an essential feature of the model that there should be a continuous slow relative movement (a self-diffusion or slow thermal agitation) between elements of the structure—molecules or crystallites—in order to maintain the distribution of wells which will be postulated. The rapid thermal agitation of individual atoms, or parts of large molecules, is to be regarded as a kind of "blur" superposed upon the wells under consideration. There appears no difficulty in this; movement of the kind unquestionably exists in viscous liquids, and presumably persists, on a lengthening time scale, in any material, however viscous. The exception is, of course, the perfect crystal, and some highly crystalline substances (*e.g.* quartz and mica) show the same constancy of loss angle as the amorphous glasses. The author suggests that in this case the temporary wells exist at planes of imperfection in the crystal structure, the perfect crystallites being in slow thermal agitation with respect to one another. It is perhaps significant that the loss angle of highly crystalline (non-polar) substances is usually very small, and (particularly in the case of mica) extremely variable as between visually similar specimens. The author suggests that this arises from the variable perfection of the crystals, leading to the existence of more or fewer temporary wells.

We now postulate that the number of these wells (which we hereafter name "thermal" wells) is distributed in depth according to

$$\delta N \propto e^{-v/kT}\, \delta v/kT \quad . \quad . \quad . \quad . \quad (14)$$

and justify this on the grounds that:

(*a*) The formation of a well deep enough to be of interest to us is a "rare" event, and therefore of exponential distribution.

(*b*) The energy involved in the wells is part of the energy of thermal agitation, and therefore in equilibrium with it. At least the deeper wells should thus have an approximately exponential distribution. Any term other than the exponential one will have only a logarithmic influence on the final result, and may be neglected.

(*c*) The assumption leads directly to the observed constant loss angle.

We further postulate that the viscosity (of an amorphous substance) or the mass of the crystallites in other substances, is sufficient that the wells, though fluctuating over long periods in such a way as to retain their exponential distribution, are yet of sufficient duration to be regarded as entities over a time large compared with the longest periodic time of measurement.

We now attribute the energy loss to oscillations of a polar or ionic element between a permanent deep well of the molecular or crystal structure, and a temporarily adjacent "thermal" well. It is assumed that oscillations between two similar wells (both permanent or both "thermal") are excluded; in the one case by the permanent wells being so deep that oscillations between them give a relaxation time in the "DC" range, and in the other by the fact that "thermal" wells of significant depth are too sparsely distributed for exchanges between them to contribute appreciable loss. This assumption is essential, since exchanges between the permanent wells would give a single time constant with a resultant "Debye" peak (it is possible that the usual rise of loss angle toward power frequencies may be connected with this). On the other hand, exchanges between members of an exponential population of "thermal" wells would give a loss angle rising linearly with frequency over almost the whole range of time constants present. The proof of this statement, which is laborious, must be omitted. The beginning of this type of loss may perhaps be seen in the rise of loss angle which usually occurs at very high frequencies, where the more numerous shallow wells would become important.

The precise nature of the element which oscillates is immaterial to our model. It may in different cases be a free ion, a free dipole arising from an impurity, a dipole forming part of the structure which happens at the moment to be suitably situated near a "thermal" well, or possibly even a charge distribution within the molecule. The distance (or angle) between wells may also be neglected in a first approximation, since such considerations will cause only logarithmic changes in comparison with the exponential distribution in number.

We proceed to show that exchanges between a number of permanent, all equally deep, wells and an exponential population of "thermal" wells, leads immediately to a constant loss angle. Define:

$\tau = A e^{v/kT}$ = the time constant of the "thermal" wells;
τ_0 = the time constant of the permanent wells;
$\tau_1 = 2\tau\tau_0/(\tau + \tau_0)$ = the time constant of the pair (see eq. (12)).

Using equations (4) (13) and (14) we write for the loss factor

$$e'' = B\int_A^\infty \frac{\omega\tau_1}{1 + \omega^2\tau_1^2} \cdot \frac{\tau\tau_0}{kT(\tau + \tau_0)^2} \cdot \frac{e^{-v/kT}}{kT} \cdot \frac{dv}{d\tau}\, d\tau \quad . \quad (15)$$

where the various terms have been kept separate to display their meaning, and where B is an omnibus constant including the number of deep wells, the distribution constant of the "thermal" wells, and the moment of the oscillating elements. Space precludes a detailed solution of the integral, but it can be reduced to

$$e'' = \frac{2B\omega A\tau_0^2}{kT}\int_A^\infty \frac{d\tau}{\tau^2(\tau + \tau_0)\{(1 + \tau_0/\tau)^2 + (2\omega\tau_0)^2\}} \quad . \quad (16)$$

and solved by substituting $z = 1 + \tau_0/\tau$. Then, subject to the condition $\omega\tau_0 \gg 1 \gg \omega A$ the result reduces to

$$e'' = \frac{(\pi - 1)AB}{2\tau_0 kT} \quad . \quad . \quad . \quad . \quad (17)$$

This expression is independent of frequency, subject to the condition stated, that the frequency shall be high compared with the time constant of the permanent wells, and low compared with the shortest time constant (A) in the "thermal" distribution. As the former time constants probably lie (from absorption current values) in the range from seconds to hours, while the latter (A) may be in the infra-red, there is no difficulty in satisfying the condition.

The result is too successful, in the sense that it gives an exactly constant loss angle (provided e' is only a small fraction of the total dielectric constant) whereas real materials of the type show an increase of loss angle at both high and low frequencies. However, the theory as given is doubtless too crude to reproduce minor details. Moreover, we have neglected exchanges between wells of similar type, and these, as mentioned in the description of the model, will lead to a Debye peak at very low frequencies and a linear rise at extremely high frequencies. It is possible, therefore, that the model explains more than is contained in the limited solution given.

The variation of loss angle with temperature predicted by equation (17) is of some interest, though the author is doubtful whether the accuracy of the theory is sufficient to allow this prediction to be made. Equation (17) may be put in the form

$$e'' = (\pi - 1)B/2kTe^{v_0/kT} \quad . \quad . \quad . \quad (18)$$

and it is then seen that there should be a maximum of the loss angle at a temperature such that $v_0/kT = 1$. This, if confirmed, would permit calculation of v_0, the depth of the permanent wells. For materials of the type in question, which all have very small loss angles, more accurate data at several temperatures would be required to enable the calculation to be made.

There is evidently a possible application of the model to the analogous problem of the internal frictional loss angle resulting from an alternating mechanical stress, which also is known to vary slowly with frequency.

In conclusion, the author wishes to express his thanks to the Director of the E.R.A. for facilities to write and permission to publish this paper. The author is also indebted to several colleagues for many helpful discussions.

British Electrical and Allied Industries Research Assn.,
Greenford,
Middlesex.

ON THE FREQUENCY DEPENDENCE OF THE DIELECTRIC CONSTANT IN DIPOLAR SOLIDS.

BY R. A. SACK.

Received 14th March, 1946.

1. Introduction.

A theoretical derivation of the frequency dependence of the dielectric constant ϵ of a dipolar substance has been given by Debye[1] on the basis of the Clausius-Mosotti assumption of the internal field. His result is expressed by the equation

$$\epsilon = \epsilon' - i\epsilon'' = \epsilon_0 + \frac{\epsilon_1 - \epsilon_0}{1 + i\omega\tau} \quad . \quad . \quad . \quad (1)$$

where ϵ_1 and ϵ_0 denote the limiting values of ϵ for very low and very high frequencies respectively, ω is the angular frequency of the applied field and τ the time of relaxation. The real part ϵ' represents the effective dielectric constant, whereas ϵ'' gives a measure of the dielectric loss. For the loss angle ϕ formula (1) yields

$$\tan\phi = \frac{\epsilon''}{\epsilon'} = \frac{(\epsilon_1 - \epsilon_0)\omega\tau}{\epsilon_1 + \epsilon_0\omega^2\tau^2} = \frac{1}{\epsilon'}\sqrt{(\epsilon_1 - \epsilon')(\epsilon' - \epsilon_0)}. \quad . \quad . \quad (2)$$

The formulæ (1) and (2) apply equally to liquids where each dipole can rotate under the action of orienting and frictional forces, and to solids where each dipole can only take up a finite number of different positions, the transition from one direction to another being treated as discontinuous. It is found, however, that for most dipolar substances the equations (1) and (2) are not in quantitative agreement with experimental results, and the theory has to be improved. The modifications of the Clausius-Mosotti hypothesis due to Onsager[2] and Debye[3] in his later papers, though leading to different values of ϵ_1 and τ, leave the formulæ (1) and (2) unchanged.

In the present paper we enquire whether the experimental deviations from (1) and (2) can be accounted for if the long range forces inherent in the Clausius-Mosotti assumption are replaced by a short range interaction of the type first considered by Ising.[4] Ising assumes that each dipole can only point in either of two opposite directions, and that there is a definite energy difference according to whether two neighbouring dipoles are parallel or anti-parallel. He calculates the total induced moment of a linear chain in a constant field by means of summations of infinite series and finds that such a short range interaction does not give rise, in the one-dimensional model, to ferromagnetism or its electrical analogue; the general shape of the curve showing polarisation against field is the same as in the Clausius-Mosotti case, only the slope is altered. Kramers and Wannier[5] have derived the same result by a different method involving the eigenvalues of matrices.

To extend the calculations to variable fields we have first to reconsider the case of a constant field and take account of the fact that the equilibrium is not static, but statistical. Each dipole is capable of reversing its direction, and this process will occur with a definite probability, depending on the directions of the dipole and its neighbours. These transition probabilities will be calculated in § 2, but it is found that consideration of any dipole in relation to its neighbours is not sufficient to deal with the case of variable fields. In § 3, therefore, the chain will be considered as a whole, and an equation derived for its behaviour in an alternating field; this is found to be identical with the Debye formula (1). In § 4 this negative result is briefly discussed and suggestions are made for further theoretical research in this line.

---

[1] Debye, *Polare Molekeln* (Leipzig), 1929.
[2] Onsager, *J. Amer. Chem. Soc.*, 1936, **58**, 1886.
[3] Debye, *Physikal. Z.*, 1935, **36**, 100, 193.
[4] Ising, *Z. Physik.*, 1925, **31**, 253.
[5] Kramers and Wannier, *Physic. Rev.*, 1941, **60**, 252.

2. Transition Probabilities.

Consider a linear chain of $2n$ dipoles of moment μ which can take up either of the two positions pointing in the direction of the chain. We shall refer to these directions as the + and − directions, and accordingly speak of + or − dipoles. Let the interaction energy between two neighbouring parallel dipoles be $-J$, between antiparallel neighbours, $+J$. A configuration having $(n+m)$ dipoles in the + direction will have a total electric moment $2m\mu$, and in an electric field of magnitude E making an angle θ with the + direction, the total energy of the chain will be given by

$$(4s-2n)J - 2m\mu E\cos\theta. \quad . \quad . \quad . \quad . \quad (3)$$

Here s denotes the number of + dipoles immediately followed by − dipoles. The number of possible configurations of external moment $2m\mu$ and internal energy $(4s-2n)J$ is

$$P(s,m) = 2\binom{n+m}{s}\binom{n-m}{s}. \quad . \quad . \quad . \quad . \quad (4)$$

(Small correction terms due to the end links of the chain have been neglected in (3) and (4).) The behaviour of the chain under the action of a constant field at absolute temperature T is determined by the partition function

$$Z = \sum_{s,m} P(s,m)\exp\left[(2n-4s)K + 2m\alpha\right] \quad . \quad . \quad (5)$$

where $$K = J/kT, \quad \alpha = \mu E\cos\theta/kT. \quad . \quad . \quad . \quad (6)$$

By means of summations of infinite series Ising finds for large n

$$Z = [e^K\cosh\alpha + \sqrt{(e^{2K}\sinh^2\alpha + e^{-2K})}]^{2n}. \quad . \quad . \quad (7)$$

The total moment of the chain is obtained by differentiation

$$M = \mu\frac{\partial\log Z}{\partial\alpha} = \frac{2n\mu e^K\sinh\alpha}{(e^{2K}\sinh^2\alpha + e^{-2K})^{\frac{1}{2}}} \approx 2n\mu\alpha e^{2K} = 2n\frac{\mu^2 E\cos\theta}{kT}e^{2J/kT}. \quad (8)$$

The latter approximation is valid if

$$\alpha \ll 1; \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

we are not interested in cases where this condition is not fulfilled.

The probability of finding any ν successive dipoles in the interior of the chain in any particular configuration can either be evaluated by Ising's summation method, by counting separately groups of 1, 2, 3 . . ., $\nu - 2$ successive + dipoles adjoining − dipoles on both sides; or, in Kramers' and Wannier's method, by forming the product of the eigenvectors of ν successive links corresponding to an addition of one dipole to the left or to the right (cf. [5]). For $\nu = 4$ the probabilities are found to be:

$$\left.\begin{array}{ll}A_4 = (++++) = & e^{3K} + \alpha(e^{5K} + 3e^{3K}) \\ B_4 = (+++-) = (-+++) = & e^{K} + 2\alpha e^{K} \\ C_4 = (++-+) = (+-++) = & e^{-K} + \alpha(e^{K} + e^{-K}) \\ D_4 = (++--) = (--++) = & e^{K} \\ E_4 = (+-+-) = (-+-+) = & e^{-3K} \\ F_4 = (+--+) = & e^{-K} + \alpha(e^{K} - e^{-K})\end{array}\right\} \times \frac{1}{16\cosh^3 K} \qquad (10)$$

$G_4 = (+---)$, $H_4 = (-++-)$, $I_1 = (-+--)$, $K_4 = (----)$ are obtained by replacing α by $-\alpha$ in the expressions for B_4, F_1, C_4, and A_4 respectively.

For three successive links the probabilities are

$$\left.\begin{array}{l}A_3 = (+++) = A_4 + B_4;\ B_3 = (++-) = (-++) = B_4 + H_4 \\ C_3 = (+-+) = C_1 + E_4;\ D_3 = (+--) = (--+) = G_4 + F_4 \\ E_3 = (-+-) = E_4 + I_4;\ F_3 = (---) = G_4 + K_4\end{array}\right\} \qquad (11)$$

If we now consider the possibility of a dipole suddenly reversing its direction, the probability of such a process taking place per unit time depends on the original position of the dipole and its two neighbours. According to the six configurations (11) we have to calculate six transition probabilities:

p_{AC} for a transition from $+++$ to $+-+$
p_{BD} ,, ,, ,, $++-$ to $+--$
p_{EF} ,, ,, ,, $-+-$ to $---$

and correspondingly p_{CA}, p_{DB} and p_{FE}.

As each reversal affects three triplets it appears that we have to determine the quantities A_5 etc., *e.g.* for a transition from $++---+$ to $+++--+$. It can be shown, however, the configurations are grouped in twos, thus being reduced to A_1, etc. We find

$$\left.\begin{array}{l}dA_3/dt = -p_{AC}(3A_4 + B_4) - 2p_{BD}B_4 + p_{CA}(3C_4 + E_4) + 2p_{DB}D_4 \\ dB_3/dt = p_{AC}(A_4 - B_4) - 2p_{BD}H_4 + p_{CA}(E_1 - C_4) + 2p_{DB}I_4 \\ dC_3/dt = p_{AC}(A_4 + B_4) - 2p_{BD}C_4 - p_{CA}(E_4 + C_4) + 2p_{DB}F_4 \\ \qquad\qquad - 2p_{EF}E_4 + 2p_{FE}G_4\end{array}\right\} \qquad (12)$$

and three corresponding expressions for dD_3/dt, dE_3/dt and dF_3/dt. When equilibrium is established in a constant field, the quantities A_4, etc., are given by (10), whereas the time derivatives of A_3, etc., *i.e.* the expressions (12) must vanish. Solving (12) for the transition probabilities, p_{AC}, etc., we obtain:

$$\left.\begin{array}{ll}p_{AC} = p_1 e^{-2K}(1 - \alpha\rho); & p_{CA} = p_1 e^{2K}(1 + 2\alpha - \alpha\rho) \\ p_{BD} = p_2(1 - \alpha); & p_{DB} = p_2(1 + \alpha) \\ p_{EF} = p_1 e^{2K}(1 - 2\alpha + \alpha\rho); & p_{FE} = p_1 e^{-2K}(1 + \alpha\rho)\end{array}\right\}. \qquad (13)$$

with three arbitrary constants, p_1, p_2, and ρ.

The knowledge of the six p's is not sufficient to solve the equations (12) for variable fields; at least one of the quantities A_4, etc., enters the equations as an indeterminable quantity.

3. Variable Field.

For a non-stationary field we consider the chain as a whole, and assume that only one dipole can be reversed in each elementary process. By each reversal the number s will either be changed by one or remain unaltered, whereas m will increase or decrease by one. There will be a certain probability that the whole chain is in any one of the 2^{2n} possible configurations. We assume that we may average over all the states having the same internal energy and electric moment. If small effects due to the end links are neglected, the mean probability, $W(s, m)$, of finding the chain in a state specified by $(4s - 2n)J$ and $2m\mu$ can be expressed as

$$W(s, m) = \frac{e^{(2n-4s)K}}{(2\cosh K)^{2n}}[1 + \alpha w(s, m)] \qquad . \qquad . \qquad (14)$$

provided the condition (9) holds. The total number $P(s, m)$ of such configurations is given in (4). The total number $P_{AO}(s, m)$ of possible transitions $(++++)$ to $(+--+)$ from all the states specified by (s, m) to a state $(s + 1, m - 1)$ can be evaluated by a similar summation method as that leading to (10) and (11); the same applies to the transition possibilities P_{BD} $(++-$ to $+--)$, P_{OA}, P_{DB}, P_{EF} and P_{FE} (*cf.* notation of triplets in (11)). These numbers are found to be

$$\left.\begin{aligned} \frac{P_{AO}(s, m)}{P(s, m)} = \frac{(n + m - s)^2}{n + m}; \quad \frac{P_{BD}(s, m)}{P(s, m)} = 2\frac{s(n + m - s)}{n + m} \\ \frac{P_{OA}(s, m)}{P(s, m)} = \frac{s^2}{n - m} \end{aligned}\right\} \quad . \quad (15)$$

and correspondingly for P_{FE}, P_{DB}, P_{EF} by substituting $-m$ for m. It follows directly from the definition of the P_{AO}, etc., that

$$P_{AO}(s, m) = P_{OA}(s + 1, m - 1); \quad P_{BD}(s, m) = P_{DB}(s, m - 1), \text{ etc.} \quad (16)$$

The general equation for the probability distributions $W(s, m)$ in any field is given by

$$\left.\begin{aligned} &P(s, m) \times \frac{\mathrm{d}W(s, m)}{\mathrm{d}t} \\ &= -W(s, m)[p_{AO}P_{AO} + p_{BD}P_{BD} + p_{OA}P_{OA} + p_{DB}P_{DB} + p_{EF}P_{EF} + p_{FE}P_{FE}] \\ &+ p_{OA}P_{AO}W(s+1, m-1) + p_{DB}P_{BD}W(s, m-1) + p_{FE}P_{EF}W(s-1, m-1) \\ &+ p_{AO}P_{OA}W(s-1, m+1) + p_{BD}P_{DB}W(s, m+1) + p_{EF}P_{FE}W(s+1, m+1) \end{aligned}\right\} (17)$$

We introduce the values for p_{AO}, etc. from (13), for W from (14), for $P(s, m)$ from (4), and for $P_{AO}(s, m)$, etc. from (15). For the quantities w in (14) we are entitled to put

$$w(s \pm 1, m \pm 1) = w(s, m) \pm \partial w/\partial s \pm \partial w/\partial m \quad . \quad . \quad (18)$$

Introducing a new variable $x = s/n$ and neglecting second order terms in m/n, we find that we can put

$$w(s, m) = m \,.\, y(x); \quad . \quad . \quad . \quad . \quad (19)$$

for a simple harmonic field $\alpha = \alpha_0 e^{i\omega t}$ cancellation of common terms leads to the equation

$$\left.\begin{aligned} 2p_1[(1-x)^2e^{-2K} - x^2e^{2K}]\frac{\mathrm{d}y}{\mathrm{d}x} + y\{2p_1[x^2e^{2K} - e^{-2K}(1 - x^2)] - 4p_2x^2 - i\omega\} \\ = 4p_1[x^2e^{2K} - e^{-2K}(1 - x^2)] - 8p_2x^2 \end{aligned}\right\} (20)$$

The coefficient of $\mathrm{d}y/\mathrm{d}x$ vanishes for $x_0 = e^{-K}/(e^K + e^{-K})$. A particular solution of (20) is

$$y = \frac{2}{1 + i\omega\tau} + \Gamma_1(x - x_0) + \Gamma_2(x - x_0)^2 + \ldots \quad . \quad (21)$$

where

$$\tau = \frac{(e^{2K} + 1)^2}{4(p_2 + p_1)} \quad . \quad . \quad . \quad . \quad (22)$$

and the Γ_i represent certain coefficients which may depend on ω. The solution of the homogeneous equation (20) becomes infinite for $x = x_0$; its factor must be zero to permit a physical interpretation of the result.

The total moment of the chain at any instant is

$$M = 2\mu\frac{\sum\limits_{s,m} mP(s, m)W(s, m)}{\sum\limits_{s,m} P(s, m)W(s, m)} \quad . \quad . \quad . \quad (23)$$

or in first approximation

$$M = 2\alpha\mu \frac{\sum\limits_{s,m} P(s, m)\, \mathrm{e}^{(2n-4s)K} m^2 y(s)}{\sum\limits_{s,m} P(s, m)\, \mathrm{e}^{(2n-4s)K}} \quad . \quad . \quad . \quad (24)$$

The only factor which depends on the frequency ω is $y(s)$, the other factors depend only on the configuration of the chain. It is easily shown that for varying s the product $P(s, m)\, \mathrm{e}^{(2n-4s)K}$ has a maximum for $s = nx_0$ which becomes more pronounced as n increases. Thus for very long chains we are only interested in the value of y for $s = nx_0$ and in (24) can regard y as a constant given by the first term in the series (21). Without carrying out the calculations in detail we see immediately that

$$\left(\frac{M}{\alpha}\right)_\omega \div \left(\frac{M}{\alpha}\right)_{\omega=0} = \frac{1}{1 + i\omega\tau} \quad . \quad . \quad . \quad (25)$$

where τ is defined in (22).

By comparison with (8) we find

$$M = \frac{2n\mu^2 \cos\theta}{kT} \frac{\mathrm{e}^{2J/kT}}{1 + i\omega\tau} E. \quad . \quad . \quad . \quad (26)$$

If the material is made up of a large number of chains the directions of which are distributed at random, we have to consider the component of M along the field E which is $M \cos\theta$. Thus the average moment per dipole in the direction of the field is

$$\bar{\mu} = \frac{2\mu^2}{3kT} \frac{\mathrm{e}^{2J/kT}}{1 + i\omega\tau} E ; \quad . \quad . \quad . \quad . \quad (27)$$

here the factor $\cos^2\theta$ has been replaced by $1/3$.

If N is the number of dipoles per unit volume, we have

$$\epsilon = \epsilon_0 + 4\pi N \frac{\bar{\mu}}{E} = \epsilon_0 + \frac{8\pi\mu^2 N}{3kT} \frac{\mathrm{e}^{2J/kT}}{1 + i\omega\tau} \quad . \quad . \quad (28)$$

which is the same as Debye's formula (1).

4. Conclusions.

The final result derived from the assumptions stated in § 1, is expressed by the equation (28); the dielectric constant depends on frequency in accordance with the Debye formula (1). Ising's conclusion that the assumption of a short-range interaction in one dimension does not lead to any new features in the static case is thus found to apply also for variable fields. This result, however, does not give any indication as to the frequency dependance of the dielectric constant in the two or three dimensional cases. Such systems show the phenomenon of ferro-magnetism or ferro-electricity below a certain transition temperature (*cf. e.g.* [6, 7]); for constant fields their polarisation has recently been calculated by Fröhlich,[8] who finds that the dielectric constant has a singularity at the transition point. It would be of interest to extend these calculations to variable fields, but until this is done, it is impossible to decide whether the assumption of short range forces can lead to a dispersion formula different from Debye's.

[6] Bethe, *Proc. Roy. Soc., A*, 1935, 150, 552.
[7] Onsager, *Physic. Rev.*, 1944, 65, 117.
[8] Fröhlich, *Proc. Roy. Soc., A*, 1946, 185, 399.

Summary.

The Debye theory of dipolar dielectrics leads to a general formula expressing the dielectric constant in alternating fields in terms of its values at very low and high frequencies, a relaxation time and the frequency. For many dipolar solids this formula is not in good agreement with experimental results and it is considered whether these discrepancies can be explained by means of a short range interaction between neighbouring dipoles. The one-dimensional problem is treated quantitatively and it is found that for this case the Debye formula is still applicable. The question whether such an interaction in two or three dimensions would account for the observed discrepancies is left open.

I wish to thank Dr. H. Frohlich for suggesting the problem and for his interest and advice concerning the work in all its stages.

H. H. Wills Physical Laboratory,
University of Bristol,
Royal Fort, Bristol 8.

Dept. of Appl. Maths.,
The University,
Liverpool.

A NOTE ON THE ANALYSIS OF DIELECTRIC MEASUREMENTS.

By S. Whitehead.

Received 15*th March*, 1946.

(1) General Principles of Linear Dielectric Phenomena.

It is not always possible to make the kind of measurements ideally required. For example the most characteristic phenomena may occur at too high or too low a frequency for convenient measurement. An important distinction is between linear and non-linear phenomena and the recognition of the onset of the latter may involve sensitive tests for the former. For practical purposes linear phenomena [1] may be regarded as those obeying or derived from a system of linear differential equations. Dielectric phenomena are usually of this type and now tend to be described in terms of an oscillation and relaxation spectrum. The oscillation spectrum may arise from electrons which, although not linear oscillators, have been shown to be capable of description by a distribution of equivalent linear oscillators; [2] there may be lattice distortion which is usually thought to be linear to a first approximation; there may be a rotation spectrum from groups with a resultant electric moment, a problem originally treated by Lorentz and recently corrected by Van Vleck and Weisskopf [3*] but agreed to be linear provided the electrical energy is much less than the thermal energy. This "collision-broadened band theory" also applies to dipoles in liquids with the natural frequency made zero to give a relaxation as opposed to an oscillation spectrum. Since the interaction with the structure of dipoles in solids yields a potential barrier with a certain rate of escape, the phenomena are relaxations but are also linear[4] to a first approximation even in Fröhlich's "ferromagnetic" theory[5] where the interaction between dipoles gives the reorientation of "domains": the energy differences caused by the electric field are usually small. Lastly, ionic effects are also linear, if stable, provided certain phenomena are excluded such as the limitation of ionic motion by barriers,[6] internal discharges in voids or other discontinuities,[7] ionisation in moist dielectrics,[8] field-induced escape.[9] The so-called electronic oscillation spectrum does not usually fall lower in frequency than the infra-red, the "rotation" spectrum may extend as low as 10^{10} c./sec. in certain cases,[2] the dipolar relaxation spectrum is usually comprised between 10^3 and 10^9 c./sec. while the ionic relaxation spectrum is below 10^3 c./sec.

---

[1] See Whitehead, *J. Sci. Inst.*, 1944, 21, 73. *E.R.A. Rep.* L/T136 (1943) and *Rep*, M/T80 (1944); Gardner and Barnes: *Transients in Linear Systems*, Vol. I (New York, John Wiley and Sons, 1942); Bayard: *Rev. Gen. Electr.*, 1935, 37, 659.

[2] Fröhlich, *E.R.A. Rep.* L T. 157 and 163, and paper in course of publication.

[3] Van Vleck and Weisskopf: *Rev. Mod. Physics*, 1945, 17, 227.

[*] Also by Fröhlich, see reference [2].

If the permittivity and loss angle of the dielectric are represented as a complex dielectric constant $\epsilon_1 - \mathrm{j}\epsilon_2$, then the admittance of any specimen is the product of a geometrical shape factor and $\mathrm{j}\,\omega(\epsilon_1 - \mathrm{j}\epsilon_2)$ while the loss angle or power factor is ϵ_2/ϵ_1. If the hypothesis of linearity is correct $\mathrm{j}\omega(\epsilon_1 - \mathrm{j}\epsilon_2)$ (if $\omega = 2\pi \times$ frequency) is a certain type of complex function of ω. It is the aim of the present note to mention some ways in which this kind of analysis may be helpful. In the past the analysis has mainly been used in circuit theory where functions such as $\mathrm{j}\,\omega(\epsilon_1 - \mathrm{j}\epsilon_2)$ are known as Kirchhoff functions and, in fact, a linear dielectric theory can be represented as a model electrical circuit as has sometimes been done. The basic formulæ generally used are as follows:

$$\omega\epsilon_2 = (1/\pi)\int_{-\infty}^{\infty} u\epsilon_1(u)\mathrm{d}u/(\omega - u), \quad \omega\epsilon_1 = -(1/\pi)\int_{-\infty}^{\infty} u\epsilon_2(u)\mathrm{d}u/(\omega - u) \qquad (1)$$

If $W(t)$ is the current due to unit constant voltage applied at $t = 0$

$$W(t) = \int_{-\infty}^{\infty}(\epsilon_1 - \mathrm{j}\epsilon_2)e^{\mathrm{j}\omega t}\mathrm{d}\omega/2\pi \quad . \quad . \quad . \quad (2)$$

If the applied voltage is $E(t)$ the current i is

$$i = \int_0^t W(t - u)\{\mathrm{d}E(u)/\mathrm{d}u\}\mathrm{d}u \quad . \quad . \quad . \quad (3)$$

Equation (1) is identical with that used by Gross.[10]* Few of the infinite number of skew-reciprocal functions ϵ_1 and ϵ_2†, are used in dielectric theory. The collision-broadened band, for example, yields, neglecting other contributions,

$$\left.\begin{aligned} \epsilon_1 &= A\left\{\frac{1 + \omega_0(\omega + \omega_0)T^2}{1 + (\omega + \omega_0)T^2} + \frac{1 - \omega_0(\omega - \omega_0)T^2}{1 + (\omega - \omega_0)^2T^2}\right\} \\ \epsilon_2 &= A\left\{\frac{\omega T}{1 + (\omega + \omega_0)^2T^2} + \frac{\omega T}{1 + (\omega - \omega_0)^2T^2}\right\} \end{aligned}\right\} \quad . \quad (5)$$

where A is a constant independent of ω and T, ω_0 is the natural frequency of the polar group, T is the decay time constant or relaxation time of the oscillator. Putting $\omega_0 = 0$, the dipole in an amorphous liquid is obtained for which

$$\epsilon_1 = A/(1 + \omega^2T^2), \quad \epsilon_2 = A\omega T/1 + \omega^2T^2). \quad . \quad . \quad (6)$$

---

[4] Fröhlich, *Proc. Roy. Soc.*, *A*, 1944, **182**, 388; *J.I.E.E.*, 1944, **91**, I, 456.

[5] Fröhlich, see reference [4] and *E.R.A. Rep.* L/T147.

[6] Garton, *J.I.E.E.*, 1941, 88, III, 23; 1941, II, 103.

[7] Austen and Whitehead, *J.I.E.E.*, 1941, 88, III, 18; 1941, II, 88. Austen and Hackett, *ibid.*, 1944, **91**, I, 298.

[8] Whitehead and Nethercot, *Proc. Physic. Soc.*, 1935, **47**, 974.

[9] Thomas and Griffith, *J.I.E.E.*, 1942, 89, I, 487; Zener, *Proc. Roy. Soc.*, *A*, 1942, **145**, 523; Fröhlich, *E.R.A. Rep.* L/T153, in course of publication.

[10] Gross, *Physic. Rev.*, 1941, **59**, 748; 1940, **57**, 57.

* Capacitance $= C = (2/\pi)\int_0^{\infty} G(u)\mathrm{d}u/(u^2 - \omega^2)$, $G(\omega) =$ conductance . (4)

† If $z = \omega + \mathrm{j}x$ and $\phi(z)$ is any analytic function subject to certain restrictions and if $\phi(z) = U(\omega, x) + \mathrm{j}V(\omega, x)$ then $\omega\epsilon_1 = U(\omega, 0)$ and $\omega\epsilon_2 = -V(\omega, 0)$ are possible values for ϵ_1 and ϵ_2.

If, as is roughly true for mica, the loss angle δ varies little up to a frequency $\omega_0/2\pi$ and is small thereafter then ϵ_2 is $\epsilon_0 \tan \delta$ as ϵ_1 will usually differ only slightly from the constant portion ϵ_0 of the permittivity, whence

$$\pi(\epsilon_1 - \epsilon_0) = \epsilon_0 \tan \delta \log \{(\omega_0/\omega)^2 - 1\} \approx 2\epsilon_0 \tan \delta \log (\omega_0/\omega)\ ^{11} \qquad (7)$$

However (4) and (5) correspond in circuit theory to a tuned circuit and a capacitance-resistance circuit respectively. If we add to these the cases of an ideal dielectric where ϵ_1 is constant and ϵ_2 is zero and a semi-conductor in which ϵ_2 is constant and ϵ_1 indefinite, then the circuit elements of a capacitor and resistor are included. Since a circuit with any desired properties can, within certain limitations, be built up from these elements, any dielectric properties which can be described as linear can also be explained by suitably distributed oscillation and relaxation spectra, ideal and semi-conducting components. This, of course, is only true in so far as variation with frequency or type of electric field is concerned, which explains why theories based upon this kind of evidence are inherently uncertain and also why variation with temperature has proved such an important factor. Thus also the scope of the present note is essentially limited in the manner indicated in the opening paragraph.

The absorption current corresponding to (4) comprises a damped oscillation of very high frequency and unlikely to be of interest. The current which corresponds to (5) is a simple exponential decay $e^{-t/T}$. The current associated with (6) is of the order of $1/t$ for times much less than $1/\omega_0$; the formulæ are of no physical significance in the regions of times or frequencies corresponding to ω_0.

Another set of functions in extensive use for the representation of experimental results [12] is

$$\epsilon_1 = \beta\omega^{m-1}(-m)!\cos \tfrac{1}{2}(1-m)\pi,\ \epsilon_2 = \beta\omega^{m-1}\pi/2(m-1)!\cos \tfrac{1}{2}(1-m)\pi$$
$$W(t) = \beta t^{-m}, \quad 0 < m < 1 \qquad (8)$$

Equation (1) can be put in the form

$$\epsilon_1 = -(2/\pi)\int_0^\infty u\epsilon_2\, du/(\omega^2 - u^2).$$

If $\epsilon_1(0)$ is the static dielectric constant then

$$\epsilon_1(0) = (2/\pi)\int_0^\infty \epsilon_2 du/u = (2/\pi)\int_0^\infty \epsilon_2(\omega)\, d\log\omega \quad . \quad . \quad (9)$$

(2) Dielectric Constant and Loss Angle.

The ionic, dipolar or oscillatory component which is to be studied may only make a small contribution to the dielectric constant but may form the major proportion of the loss angle. For example it is possible to observe with certainty changes of $\tan \delta$ to 10^{-6} but a change of 10^{-6} in dielectric constant would, in general, be impossible to detect. For that reason inference of the parameters including the variation of dielectric constant from $\tan \delta$ variations is often attractive. If both dielectric constant and loss angle are compared it is usually only necessary to assume one parameter $\epsilon(\infty)$ for a quantitative check. This is important when only one parameter, usually the relaxation time T, depends upon temperature in an unknown fashion. In any simple relaxation, as with

---

[11] See application by Garton, *Capacitors used for Measurement Purposes*, read, but not yet published by I.E.E.

[12] Hartshorn, *J.I.E.E.*, 1926, 64, 1152; 1932, 70, 417; Hartshorn and Rushton, *ibid.*, 1938, 83, 315; Hartshorn and Ward, *ibid.*, 1935, 77, 723. Hartshorn, Rushton and Megson, *Proc. Physic. Soc.*, 1940, 52, 796.

dipoles, T is eliminated in the comparison. Garton[13] showed the inadequacy for solid resins of the Debye theory, and also that of Fröhlich which gives identical results if $\epsilon(\infty)$ is open to choice, by showing that both the tan δ and ϵ curves could not be fitted simultaneously. In his theory of a distribution of relaxation times, where the corresponding groups of dipoles are released at various temperatures, he determines the distribution from one curve and shows that it agrees with the other. Although his analysis only requires an assumption of the "relaxation" Hilbert pair of equation (6), yet any other linear hypothesis could have been applied, so that the general conclusion should be possible that his distribution theory could be made to agree with any type of linear component.

Numerical integration of equation (9) to obtain the static dielectric constant is attractive but difficult to use because the integrand has really a pole at the origin ($\omega = 0$) where measurements are most uncertain. Gross[10] has used equation (4) numerically upon dielectric waxes where the variation of tan δ is slow and regular. The form given in equation (8) is susceptible to a simple approximate treatment. As already shown the different types of component are associated with widely different time-constants or frequencies, thus over each range of frequency it can be assumed that there is only one pair of values β and m for the group of components in that region. Accordingly, if ϵ or tan δ is plotted on double logarithmic paper the slope of the high ω end will give m_1 and a trial value for β_1, *i.e.* a more or less asymptotic straight line. The differences from this line can be replotted and the process repeated for m_2 and β_2 and so on. It is more usual to try to express ϵ or tan δ as a series of simple functions such as those already given for relaxation and oscillation terms, equations (5) and (6), and then to transform. A considerable body of literature exists upon this method in circuit theory since the inference of the reactance function from the resistance function of a network impedance is identical with the dielectric problem (see for example Lee[14]). Another method notes that a series of forms such as equations (5) and (6), *i.e.* $\Sigma_k \phi_k/\{1 + (\omega \pm \omega_k)^2T^2\}$, is also an expression for the potential at the point $(\omega, 0)$ on the axis of abscissæ due to point charges situated at the points $(\pm \omega_k, \pm 1/T_k)$. For the relaxation spectrum the point charges are all on the axis of ordinates. If ϵ_1 or ϵ_2 were more or less constant up to ω_0, then the segment 0 to ω_0 of the abscissæ would be an equipotential line, whence it follows that there would be a concentration of point charges near ω_0 and that such a variation of dielectric properties if any such discontinuity is shown implies the presence of oscillatory components since relaxation components on the axis of ordinates could only give a rough approximation to an equipotential other than zero.* Both from this analogy and by inspection of equation (1) it can be seen that correlation between the variations of ϵ_1 and ϵ_2 is only to be expected over the same ranges of frequency.

(3) Absorption Currents.

Suppose $W(t)$ has a constant part g and a variable part $\phi(t)$, and suppose the field E has an initial value E_0. If the time variation of E is much less than that of ϕ then equation (2) gives

$$i \approx gE + E_0\phi(t) + (\mathrm{d}E/\mathrm{d}t)\int_0^t \phi(u)\mathrm{d}u \approx E_0\phi(t) + gE + \epsilon^1(\mathrm{d}E/\mathrm{d}t) \qquad (10)$$

This defines an anomalous conductance g and anomalous dielectric constant ϵ_1^1. It has been assumed implicitly that most of the rapid variation of ϕ is complete by t. Suppose the interval for completion is u_0 then

---

[13] Garton, *J.I.E.E.*, 1939, 85, 625. [14] Lee, *J. Math. Physics*, 1932, 11, 83.

* This is also shown by the absorption current $W(t)$ for the case where ϵ_1 has a constant value ϵ_0 up to ω_0 and is thereafter a lower value ϵ_∞. $W(t)$ is then $\{2(\epsilon_0 - \epsilon_\infty)/\pi t\}\mathrm{Si}\ \omega_0 t$ which has a decaying oscillatory component.

The conductance g is related by an operational equation with the imaginary dielectric constant so that equivalent values of the complex dielectric constant can be deduced from the absorption current with respect to various slowly varying types of voltage. This is of use in problems such as the redistribution of charge between dielectric layers of different properties and an elementary instance is the familiar term " capacitance for a given time of charge."

Can a value for u_0 be found which will give the correct values of ϵ_1 and ϵ_2 for sinusoidal fields ? It follows from equation (2) that u_0 must then be a solution of

$$\int_0^{u_0}\phi(u)\,\mathrm{d}u = \int_0^{\infty}\phi(u)\cos\omega u\,\mathrm{d}u \text{ and } \phi(u_0) = \int_0^{\infty}\omega\phi(u)\sin\omega u\,\mathrm{d}u. \quad (11)$$

If $\phi(u)$ is of form $e^{-t/T}$ the two equations have the same solution

$$u_0 = T\log\{1 + (\omega T)^{-2}\} \quad . \quad . \quad . \quad . \quad (12)$$

Thus for any given frequency ϵ_1 and ϵ_2 can be defined by the absorptive charge and absorption current after a certain interval of charge which varies with the frequency. If, however, the absorption current were to vary as t^{-m}, then to obtain the correct value for ϵ_1, ωu_0 would have to be $\{(1 - m)!\cos\frac{1}{2}(1 - m)\pi\}^{1/1-m}$ while to obtain the correct value for ϵ_2, ωu_0 would have to be $\{(1 - m)!\sin\frac{1}{2}(1 - m)\pi/1 - m\}^{-1/m}$. Although the interval u_0 is then always a constant fraction of a cycle the fraction is different for the permittivity and loss angle. Thus these " absorption " definitions can only be used in an approximate fashion in dealing with sinusoidal fields. Nevertheless, the absorption charge and current can be obtained conveniently from equation (2) by measurements of ϵ and $\tan\delta$ over the range of frequencies corresponding to intervals which are too short for convenient measurements by applying a field for a known time. Fig. 1 shows the variation of capacitance or dielectric constant with time of charge over 6 decades deduced from A.C. measurements down to 50 cyc./sec. and D.C. measurements from about 0·001 to 10 sec. The A.C. capacitance and loss angle results afforded an internal check and the results show the characteristic features of the short time relaxation terms which are clear in the − 11° C. curve but have moved to much shorter times in the higher temperature curve and the long-time relaxation terms apparent at the higher temperatures but outside the range at − 11° C. These latter terms are probably ionic as already noted and are often forgotten in descriptions of dielectric properties. For example, these terms are of the same order in mica as in some lower grade dielectrics whereas they are much smaller in polythene and poly-styrene.[11] In Fig. 1 the h.f. capacitance (here called C_0) has been taken as that at −11° C. for a charging time of about 1 microsecond, an arbitrary assumption which excludes all oscillation terms.

The preceding analysis was made partly by numerical integration of the two equations implicit in equation (2) and also by representing the results by the sum of one or two of the simple related functions given in section (1). The latter is the usual method and for the purposes of physical significance the absorption current is usually analysed into a sum of exponential relaxation terms since the oscillation spectrum has not yet been studied. J. B. Whitehead[15] has made considerable use of this analysis

$$g = W(u_0),\ \epsilon_1^1 = \int_0^{u_0} W(u) - W(u_0)\mathrm{d}u.$$

[15] J. B. Whitehead and Eager, *J. App. Physics*, 1942, 13, 43. *Essays in Rheology*, Essay III : *The Time Variation of Stress and Strain*, Isaac Pitman, 1946. Symposium prepared by British Rheologists' Club. The section relating to exponential analysis will be amplified in a forthcoming paper by Elleman and Lethersich ; Simha, *J. Physic. Chem.*, 1943, 4, 348.

by the method of successive graphical approximation (this time on semi-logarithmic paper) as already described in finding a sum of ω^m terms. The same problem arises in the same way in rheology and three methods of exponential analysis have been used: one by finding the "best fit" or regression line after the exponential terms have been reduced to simple linear terms, a method which requires usually too high an experimental accuracy; the successive approximation method already mentioned; a plot of $W(t)$ against $\log t$ in which a point of inflexion should occur in the region where t is equal to any relaxation time, but this method is rather too inaccurate.[15] Of these three methods the second has been the most successful. If $W(t)$ is the Laplace transform of $U(p)$ then

$$W(t) = \int_0^\infty U(p) e^{-pt} \, dp \quad . \quad . \quad . \quad . \quad (13)$$

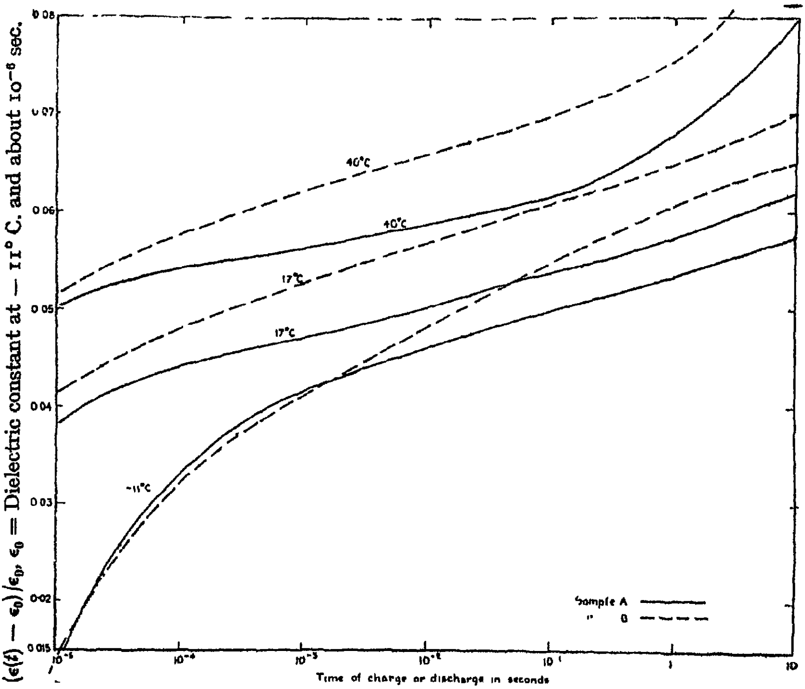


FIG. 1.—Variation of apparent dielectric constant with time at constant field of 2 samples of oiled paper.

Thus the partition function for a continuous exponential distribution with respect to the relaxation time T is

$$(1/2\pi j T^2) \int_{k-j\infty}^{k+j\infty} W(t) \, e^{t/T} dt.$$

This is not very suitable for numerical computation from a curve for $W(t)$. An interesting case is when $W(t)$ is t^{-m}. The partition function is then

$$(1/2\pi j T^{m+1}) \int_{k-j\infty}^{k+j\infty} W(t) \, e^{t/T} dt$$

which is not really significant and corresponds to the difficulty in finding a physical theory which leads to a power law. However, a formal mathematical result can be obtained of a distribution $T^{1-m}((m-1)!)^{-1} e^{-t/T}(dT/T^2)$. Accordingly as m is greater or less than unity, the main relaxation times are greater or less than the times in consideration.[16] Whatever its theoretical limitations this way of presenting results is very useful experimentally as has been shown by Hartshorn and others.[12]

---

[16] Whitehead, *E.R.A. Rep.* L/T 151, 1944.

Owing to the segregation of the various oscillation and relaxation terms into groups, repeated rectangular pulses can be used to observe absorption over short time intervals and are also of specific interest in connection with the pulse properties of dielectrics which if approached from the A.C. side might require inconveniently high frequencies. The pulse properties of dielectrics form an extensive subject and here it is only possible to indicate that the formulæ already given provide in principle means of study which have not yet been explored experimentally to any extent.

An exponential pulse $e^{-t/T}$ or an exponential step $(1 - e^{-t/T})$ acting on an exponential term $e^{-t/T}$ of same time constant yield currents of $\{1 - (t/T)\}e^{-t/T}$ and $(t/T)e^{-t/T}$ respectively. Although this is a kind of analogy to resonance it is not as useful in practice as it appears because the reversal of sign in the former occurs when the current is rather low nor is the enhancement of current in the latter sufficiently marked.

If a field E_0 is applied to a dielectric from 0 to t, when the dielectric is short-circuited until t_2, then a residual charge or potential E builds up if the dielectric is then open-circuited. E is given by an integral equation of the second kind

$$0 = E_0\{W(t) - W(t - t_1)\} + \int_{t_2}^{t} W(t - u)(\mathrm{d}E/\mathrm{d}u)\mathrm{d}u \quad . \quad (14)$$

Although this method has been used in investigations, the expression is unhandy for experimental use except where $W(t)$ is of the form t^{-m} and if the short-circuit period is eliminated or substantially instantaneous when Abel's integral equation is obtained and the solution is

$$E_0 - E = (E_0/\pi) \sin m\pi \int_{t_1}^{t} \mathrm{d}u/(t - u)^{1-m}u^m \quad . \quad . \quad (15)$$

(4) Various Time-Dependent Fields.

For the fields hitherto considered—constant, sinusoidal, rectangular pulses, exponential—the dielectric properties have a specific interest which it is desired to infer from such measurements as can be made and to discriminate linear and reversible processes from the many rather unstable or adventitious effects which may be present. Fields of other forms are used in order to detect the initiation of non-linear phenomena particularly those associated with high stresses. Effects such as discharges and partial breakdowns,[7] field-induced ionisation by escape of electrons through potential barriers [9] or of bound ions, depend as regards the former on the change of field as well as on its maximum value whereas the latter depend mainly on the maximum value. All high-stress phenomena also depend upon time of application. The problem arises of applying high fields for short times to avoid undesired breakdown (such as flashover or thermal breakdown) comprising also a rapid change of field. It is proposed to describe some of the special methods used.

(4.1) Linear Rise.—If a field is applied at a constant rate α per second then a dielectric constant ϵ_1 and conductivity g both independent of frequency would yield a current $\alpha\ (\epsilon_1 + gt)$ where t is the time. It is fairly easy to deduce by comparison that when absorption is present the current no longer varies linearly with time but the current at any instant is a measure of the dielectric constant for a given time of charge while the slope of the tangent to the current-time curve yields the equivalent conductivity.[17] In practice, however, it is not easy to achieve an adequately constant rate of rise as in curve A, Fig. 2, where the unbalanced current

---

[17] Thomas and Whitehead, *E.R.A. Rep.* L/T70, 1936; Austen and Whitehead, E.R.A. Ref. L/T96, 1939.

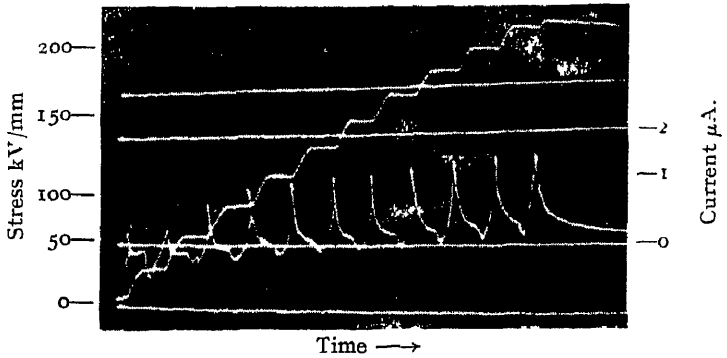


FIG. 3.—Clear ruby mica 0·083 mm. conditioned at 100° C.

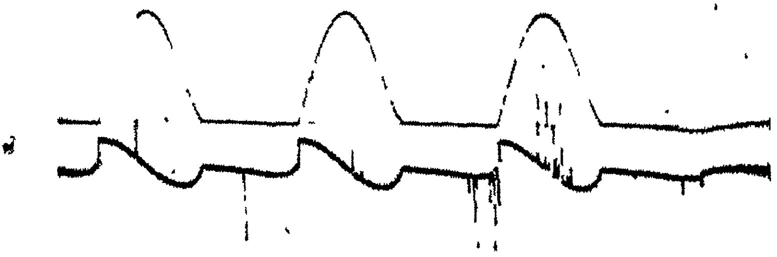


FIG. 5.—Current due to alternate half-waves showing discontinuities due to discharges (Polythene).

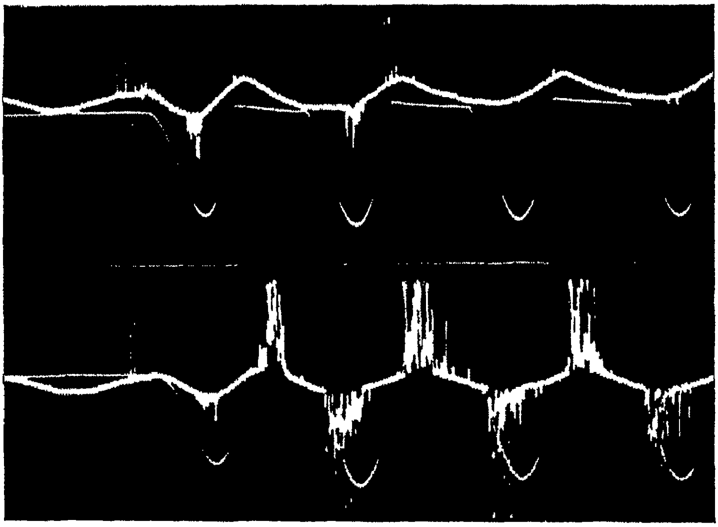


FIG. 6.—Comparison of full-wave and half-wave effects including discharge effects (Polythene).

To face page 73.]

has the wrong form on this account. Accordingly, the constant portion of the capacitance is usually balanced by a bridge network to give a curve such as B, Fig. 2, which comprises mainly the absorption current. Fig. 2, curve E, shows that the field comprised a steady rise, a constant portion and exponential decay. Comparison over the two former periods facilitates the separation of ϵ_1 and ϵ_2. The exponential decay has certain advantages already mentioned. The discontinuity when the decay starts yields a large transient in the total current which would mask absorption but is substantially balanced out in curve B. The single linear rise has not been used very much for sensitive work but in principle it satisfies the conditions of providing high fields and large changes of field rapid compared with the leakage or D.C. time constants of dielectrics.

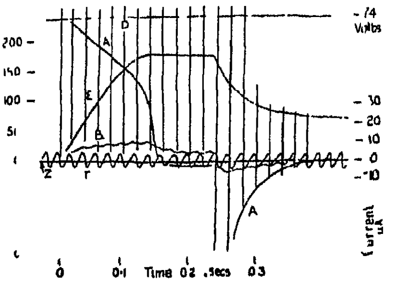


FIG. 2.—Cellulose acetate sheet, 0·09 mm., with graphite electrodes. Conditioned for 24 hours at 100° C.
Z = zero sweep.
T = time base calibration.
A = rate of change of voltage.
B = current in specimen other than geometric charging current.
D = calibrating sweep.
E = voltage applied to specimen.

(4.2) Stepped Fields.—As may be appreciated from the foregoing, the linear rise is best adapted to finding the effective dielectric constant while the decay current for a constant field is useful for the effective conductivity or loss angle. The two together form a step and during the rise the current should reach a value corresponding to the dielectric constant for a given time of charge and during the "tread" should fall to a value corresponding to the conductivity for a given time of charge. Fig. 3 shows an early record for a number of steps and the envelope of the crests should give the variation of dielectric constant with field and the envelope of the troughs (excluding the downward kicks due to imperfect balance of the constant capacitance) the variation of the conductivity. Allowance must be made for the cumulative effect of the residual absorptive charge which remains undischarged during the "tread." By this means Austen[18] showed the absence of effects of high field-strength on mica hitherto reported but if the phenomena are linear more accurate results can be obtained with A.C. and in the example mentioned Austen, using mainly A.C. tests, subsequently showed the absence of the effect of field strength on mica up to one-fifth the intrinsic electric stress. On the other hand, results over a wide range of field strength are obtained from a single record with a very short application of a unidirectional field, a considerable advantage where breakdown for sustained fields is feared. Discharges or partial breakdown appear as

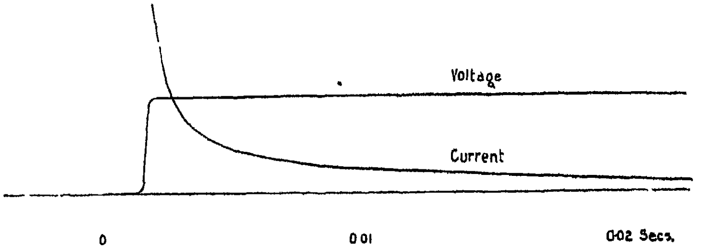


FIG. 4.—Absorption decay current in mica with constant field.

---

[18] Austen, *E.R.A. Rep.* L/T118 : *J.I.E.E.*, 1945, 92, I, 373.

discontinuities but, owing to the number of discontinuities inherently present, can be missed. If anything other than fairly rough values of loss angle or conductivity are required, care in recording is necessary. Fig. 4 illustrates the kind of curve used for measurement of absorption current, taken with a single rectangular step of voltage. It was from this kind of record that the curve of Fig. 1 and similar curves were obtained.

(4.3) **Mixed or Intermittent Fields.**—A small alternating field superimposed upon a suddenly applied constant field should yield all the quantities necessary for correlation since the mean current would follow the absorption characteristic $W(t)$ while the ripple current would yield ϵ_1 and ϵ_2. If the alternating field were superimposed on a linearly increasing field then ϵ_1 and ϵ_2 found from the ripple as before should indicate the presence of field-induced mechanisms, while the mean current could be treated as in the section on the linear rise.

In principle discharges and partial breakdown would show as discontinuities in the latter type of field. The study of this phenomenon [7] in recent years leading to a fairly quantitative theory of certain aspects has indicated the importance of its incidence with sinusoidal or rapidly varying voltages owing to the importance of field change. For this reason and a number of others including experimental convenience, consideration has been given to half waves where the unidirectional nature of the field is preserved. The current due to the sudden application of a field sin ωt is

$$i = \cos \omega t \int_0^\alpha f(u/\omega) \cos u\,du + \sin \omega t \int_0^\alpha f(u/\omega) \sin u\,du \quad . \quad (16)$$

where α is the epoch of the wave at the instant considered. The analysis of this current is not especially simple as may be seen by taking $f(t)$ as $e^{-t/T}$ when

$$i = \{\omega T/(1 + \omega^2 T^2)\}(\cos \omega t + \omega T \sin \omega t - e^{-t/T}) \quad . \quad (17)$$

Nevertheless, there are methods of dealing with such forms but this aspect has not been explored in view of the interest attaching to discharges. Fig. 5 shows a record with alternate half waves applied to a bridge network which balances the more constant portions of ϵ_1 and ϵ_2. In fact the bridge was slightly overbalanced yielding a spurious component in the absorption current, which can be allowed for, but the discontinuities due to discharges are clearly observed and easily measurable. Fig. 6 gives a comparison between half waves and a sinusoidal field. The self-extinguishing of discharges in some conditions is shown with half waves while the full wave picture shows their persistence or restoration in other conditions. The details of this type of non-linearity, as distinct from its discrimination from other phenomena, is, however, interesting outside the scope of the present note.

(5) Conclusions.

(5.1) Most dielectric theories lead to substantially linear phenomena, in fact a spectrum tending to divide into three groups: a very high frequency rotation spectrum and medium and low frequency relaxation spectra associated with dipoles or ions. Accordingly, the dielectric constant, loss angle and anomalous or absorption current are inter-related and any may be chosen for investigation according to experimental convenience.

(5.2) Comparison of all three types of measurement may permit the elimination of parameters inessential to the theory under test, particularly when it is the variation with temperature which is really crucial.

(5.3) A similar comparison often facilitates the detection of non-linear mechanisms, particularly those of a continuous type.

(5.4) Forms of field other than the usual constant and alternating types may be advantageously employed for the study both of linear and non-linear phenomena at more intense fields and more rapidly than is otherwise convenient. In particular, discontinuous non-linear phenomena can profitably be studied in this way.

Summary.

Present dielectric theories are usually linear to a first approximation and give rise to an oscillation, or rather rotation spectrum and a relaxation spectrum usually divided into two groups. The former spectrum is of very high frequency while the two latter groups are of medium and low frequency respectively. A number of general theorems apply to such phenomena, mainly studied hitherto in other connections. The dielectric constant, loss angle and absorption current are all inter-related in linear phenomena so that equivalent results can be inferred from the study of any one, while a comparison of them can eliminate various common parameters when it is desired to test a theory wherein absolute calculation of these parameters is not essential. Time dependent fields other than constant or sinusoidal can be employed for the study of certain features, particularly where it is desired to use only a unidirectional field applied for a short time. Non-linear mechanisms are of considerable importance in the industrial application of dielectrics and their detection and study can be assisted by the use of general principles, while the use of fields with special kinds of time dependance has advantages particularly with discontinuous non-linear phenomena such as discharges.

This note includes material from a number of Electrical Research Association Reports for which acknowledgement is due to the Association. The author thanks Dr. Austen and Dr. Fröhlich for permission to quote unpublished material.

British Electrical and Allied Industries Research Assn.,
15 *Savoy Street,*
London, W.C.2.

GENERAL DISCUSSION

Dr. F. C. Frank (*London*) said: Firstly, may I say that I regard Mr. Garton's paper as one of the most important before the meeting, particularly because it promises the first real understanding of the low, very flat, loss-frequency curves of the *good* solid dielectrics, for which $\tan \delta$ is practically never better than 10^{-4} and often much worse. It may either tell us how to improve them, or why we never can.

It may be of some use, to a society of physical chemists, to provide some translation of Garton's terminology. The relaxation law $e^{-t/\tau}$ is the unimolecular reaction law which the physical chemist usually writes as e^{-kt}. The expression $\tau_1 = 2\tau\tau_0/(\tau + \tau_0)$ for the time constant of two unequal wells, given at (*a*), then corresponds to $k_1 = k + k_0$, which expresses the well-known fact that when there is an equilibrium between two opposing unimolecular reactions, the rate constant for restoration of the equilibrium after a disturbance is the sum of the rate constants for the two opposing reactions. (There is a discrepancy by a factor of 2, due to a peculiarity in Garton's definitions.)

The expression given at (*b*) should correspond to $1/kT$ times the reduction factor R given in my paper. My R is regularly found to be $1 - (\overline{m}/\mu)^2$, and for a two position case, $(\overline{m}/\mu) = (n - n_0)/(n + n_0) = (k_0 - k)/(k_0 + k)$ where $(n_0/n) = (k/k_0)$. Thus $R = 4kk_0/(k + k_0)^2 = 4\tau\tau_0/(\tau + \tau_0)^2$ agreeing with Garton's result apart from an omitted factor of 4.

Now, as we have just seen, the latter expression is related to an equilibrium constant, which only depends on the difference between the two "wells", whereas the former is related to a velocity constant, which depends also on the potential barrier between them. The permanent wells are doubtless (in a crystalline substance) the "proper sites" for ions or molecules, and rightly regarded as all alike. Garton's theory implies that the potential barriers have a uniform height above these, and needs either a justification of this, or an investigation of the effect of putting an exponential or Gaussian spread on the heights of the barriers. A full theory must also include entropy terms, involving the shapes as well as the heights and depths of wells and barriers.

In the postulated distribution function for "thermal wells", it is likely that T in the exponent refers to the actual temperature in the case of a viscous substance, but probably to some temperature in the past history of a solid, and in the case of a glass to the temperature at which it actually became a glass, *i.e.* when its rate of internal readjustment ceased to keep up with the rate of cooling.

Finally, I should like to suggest that in the succeeding paragraph some such word as "transitions" would be preferable to "oscillations".

Mr. C. G. Garton (*London*) (*partly communicated*) : The expression

$$\frac{1}{\epsilon'}\,\frac{d\epsilon'}{dT} = \frac{2\tan\delta}{\pi T}\log\frac{1}{\omega\tau_0} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

(equations 21 and 22 of Messrs. Gevers and Du Pré's paper) is deduced by the authors as a consequence of a particular model of a dielectric.

The purpose of this comment is to show that it really has wider application, and that an expression of this, or very similar, type can be more simply derived without reference to any particular model.

The following differs from an alternative demonstration given verbally at the Discussion, but avoids an undetermined constant of integration which somewhat obscured the significance of the earlier method.

If any dielectric mechanism be linear with regard to the applied field, it can be represented (whatever its real nature) by

$$\epsilon' = \epsilon_\infty + A\int_{\tau_0}^{\infty}\frac{f(\tau)d\tau}{1+\omega^2\tau^2} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and

$$\epsilon'' = A\int_{\tau_0}^{\infty}\frac{\omega\tau f(\tau)d\tau}{1+\omega^2\tau^2} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The distribution function necessary to secure that tan δ remains substantially constant with respect to frequency is

$$f(\tau) = 1/\tau$$

and the solutions of (2) and (3) then become (approximately, and within the range $0 < \omega\tau_0 \ll 1$)

$$\epsilon' = \epsilon_\infty - A\log\omega\tau_0 \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$\epsilon'' = \pi A/2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

From (5)

$$A = 2\epsilon''/\pi = 2\epsilon'\tan\delta/\pi \quad . \quad . \quad . \quad . \quad . \quad (6)$$

and from this, provided $\epsilon' - \epsilon_\infty \ll \epsilon_\infty$

$$\epsilon' = \epsilon_\infty(1 - (2\tan\delta\log\omega\tau_0)/\pi) \quad . \quad . \quad . \quad . \quad (7)$$

If tan δ is the only term in (7) which is dependent upon temperature, then

$$\frac{d\epsilon'}{dT} = \frac{2\epsilon_\infty}{\pi}\log\frac{1}{\omega\tau_0}\cdot\frac{d}{dT}\tan\delta \quad . \quad . \quad . \quad . \quad . \quad (8)$$

This expression must hold for any material, whatever its mechanism of loss, which complies with the assumptions made, namely that the mechanism is linear; the loss angle is independent of frequency; the non-polar part (ϵ_∞) of the dielectric constant is independent of temperature, and the shortest time constant present is also independent of temperature. This last condition is usually implied in discussions of the distribution of relaxation times, through use of the expression $\tau = \tau_0 e^{v/T}$ to represent any constant in the distribution. The shortest time is thus always τ_0 and is independent of temperature. If the condition is not fulfilled, differentiation of (7) is more complicated, but owing to the logarithmic term the numerical value of the result is not greatly changed unless the dependence of τ_0 on T is exponential.

To transform (8) into (1) requires knowledge of the relation between tan δ and temperature, and to this extent the result is dependent upon the mechanism assumed. The next paragraph, however, will show that this relation is only logarithmically involved and may vary considerably without causing much change in the final result.

Equation (8) may be expressed in the alternative form (noting that

$$\frac{d\epsilon'}{dT}\Big/\tan\delta = \frac{2}{\pi}\log\frac{1}{\omega\tau_0}\cdot\frac{d}{dT}(\log\tan\delta) \quad . \quad . \quad (9)$$

Defining, for convenience, $G = (d/dT) \log \tan \delta$ we may assume any probable dependence of tan δ on temperature, and insert the resulting value of G in (9).

If $\tan\delta \propto T^n$ then $G = n/T$ . . . . (10)

and this (with $n = 1$) gives exactly Gevers and Du Pré's result (1).

From the writer's paper " Distribution of Relaxation Times in Dielectrics " there is some reason to expect a relation of the form

$$\tan\delta \propto T^n e^{-a/T}$$

in which case $G = (n + a/T)/T$ . . . . (11)

If the (n) term is predominant, this does not differ from (10). If, however, the (a/T) term predominates, then approximately $G = a/T^2$ and it is necessary to consider whether this may still lead to constancy of (9) for various materials. Since most data for materials of the constant loss angle type cover only a narrow temperature range (about 290° to 370° K.) the effect of the altered exponent of T is not likely to be great, while the absolute value of G is unimportant since the value of the logarithmic term in (1) is, within limits, disposable. The one remaining consideration is that the constant (a) of the material, which does not appear in (1), may vary widely between materials. Since, however, this constant contains an activation energy, exponentially related to the properties of the material, it cannot vary by more than a factor of two without either carrying the material out of the constant loss angle class, or alternatively reducing the loss angle so much that variation of ϵ_∞ predominates in the temperature coefficient. In either case, the present arguments do not apply.

It would be valuable if Messrs. Gevers and Du Pré would make available the collection of data from which their Fig. 1 has been calculated.

It is of some interest to note from (8) that if either the dielectric constant or the loss angle be independent of temperature, then (subject to the assumptions made) both must be so independent.

Dr. M. Magat (*Paris*) said : Dr. Fröhlich just pointed out that the theory developed by Professor Bauer predicts the existence of losses at frequencies above the Debye critical frequency. These losses are due to the relaxation of libration, while the usual Debye losses are due to the relaxation of rotation. This may be an explanation of the small losses found by Girard and Abadie on the high frequency side of the Debye curves for certain liquids, as well as of the large discrepancy between the refractive indices of Neoprene for 10 cm. waves, and for the sodium light as observed by Schneider, Carter, Smyth and myself. Let us remember that Maar has found for a number of plastic materials losses in the region of 0·5-0·2 mm.

Mr. Whiffen (*Oxford*) said : I notice that no recent papers have appeared which discuss the effect of dipole acceleration terms on the dispersion and loss factors at very high frequencies. From a paper by Rocard [1] it would seem that for molecules with moments of inertia of 500×10^{-40} g. cm.[8] a correction of the order of magnitude of 5 % is required at a frequency of 50,000 megacycles/sec. (6 mm. wavelength). It seems desirable to extend current theories of dielectric loss to include inertia terms ; it may be noted that Rocard assumed that Stokes' law was applicable to individual molecules.

---

[1] Rocard, *J. Phys. Radium*, 1933, 4, 247.

Dr. S. Whitehead (*London*) (*communicated*) : Part of Mr. Garton's comment on the paper by Gevers and Du Pré can be established alternatively and rather more rapidly and fundamentally, by using the results of the writer's paper, " A Note on the Analysis of Dielectric Measurements ". Now $\omega(\epsilon' + j\epsilon'')$ is a Kirchhoff function and, if $\epsilon'' \leqslant \epsilon'$, may be replaced by $\omega\epsilon' e^{j \tan \delta}$, whence from the Hilbert transform relation between tan δ and log $\omega\epsilon'$,

$$\log \omega\epsilon' = -\frac{2}{\pi}\int_0^\infty \frac{\tan \delta}{\omega - u}\, du = -\frac{2 \tan \delta}{\pi}\int_0^{l/\tau_0} \frac{du}{\omega - u}$$

if tan δ varies little with frequency up to $1/\tau_0$ and is thereafter negligible as has been assumed. If (with Mr. Garton) $0 < \omega\tau_0 \leqslant 1$,

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial T} = \frac{2 \tan \delta}{\pi}\frac{d \log \tan \delta}{dT} \log \frac{1}{\omega\tau_0},$$

is obtained by differentiation and evaluation where T = abs. temperature. Thus the formula of Gevers and Du Pré,

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial T} = \frac{2 \tan \delta}{\pi T} \log \frac{1}{\omega\tau_0}$$

means that tan δ varies little with ω up to $1/\tau_0$ but exponentially with the inverse of the temperature. These features are often encountered experimentally, and are common features of several of the theories in current discussion.

C. BREAKDOWN.

ELECTRIC BREAKDOWN OF SOLID DIELECTRICS.

By A. von Hippel.

Received 20th February, 1946.

I. Experimental Facts.

(A) Breakdown Field-strength and Pre-breakdown Currents.—A thin plate of a solid dielectric, when properly shielded in a sample holder and subjected to homogeneous d.c. fields of slowly increasing strength, goes through several stages of conduction in succession.[1] In weak fields, the current at constant voltage decreases with time; in higher fields, this decay becomes less and less pronounced and may even be replaced by an initial increase of conductivity; finally, at a critical field-strength the current rises abruptly by many orders of magnitude corresponding to the breakdown of the dielectric. This breakdown, from onset to finish, requires only about 10^{-7} seconds.[3] Somewhat below the breakdown field-strength, $E_{\max}$, the current may become unsteady. Single shots can be heard after amplification and can be measured as noise, such fluctuations growing in number and intensity with voltage and diminishing at constant voltage with time.[2, 4]

---

[1] von Hippel, *Z. Physik,* 1931, **68,** 309 ; 1932, **75,** 145 ; *Ergeb. exact. Naturw.* 1935, **14,** 79 ; *J. Appl. Physics,* 1937, **9,** 815.

[2] von Hippel, *Physic. Rev.,* 1938, **54,** 1096.

[3] Rogowski, Flegler and Tamm, *Arch. Elektrot.,* 1927, **18,** 479.

[4] Haworth and Bozorth, *Physic. Rev.,* 1932, **39,** 845 ; *Physics,* 1934, **5,** 15.

[5] von Hippel, *Z. Physik,* 1932, **75,** 145 ; *Z. Electrochem.,* 1933, **39,** 506.

[6] von Hippel, *Z. Physik,* 1934, **88,** 358 ; von Hippel and Lee, *Physic. Rev.,* 1941, **59,** 824.

[7] Buehl and von Hippel, *Physic. Rev.,* 1939, **56,** 941.

[8] Austen and Hackett, *Nature,* 1939, **143,** 637 ; Austen and Whitehead, *Proc. Roy. Soc., A,* 1940, **176,** 33.

[9] von Hippel and Maurer, *Physic. Rev.,* 1941, **59,** 820.

The breakdown field-strength of a material, measured under conditions that eliminate edge effect, is a function of structure, composition and temperature. For alkali halides it decreases systematically from Li to Rb and from F to I (Fig. 1),[5] is influenced by addition agents (Fig. 2),[6] and passes through a maximum near room temperature.[6, 7, 8] Passing from the crystalline to the amorphous state, the breakdown strength at low temperature increases and the temperature dependence in this range[9, 11] disappears, as is indicated by data on quartz (Fig. 3).[9] $E_{\max}$ is independent of the sample thickness down to 10^{-4} cm. or less, as shown in amor-

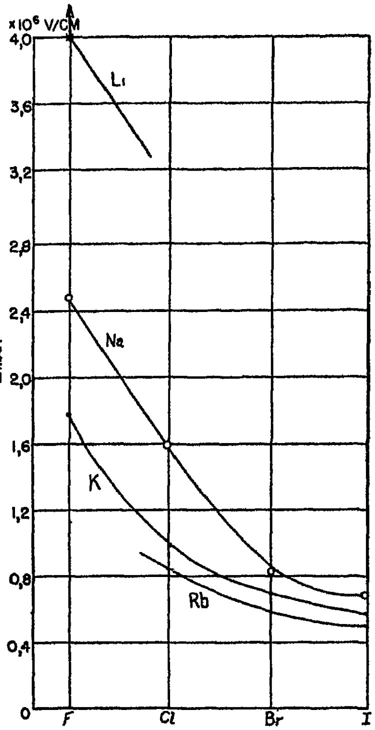


FIG. 1.—Breakdown strength of the alkali halide crystals at room temperature.

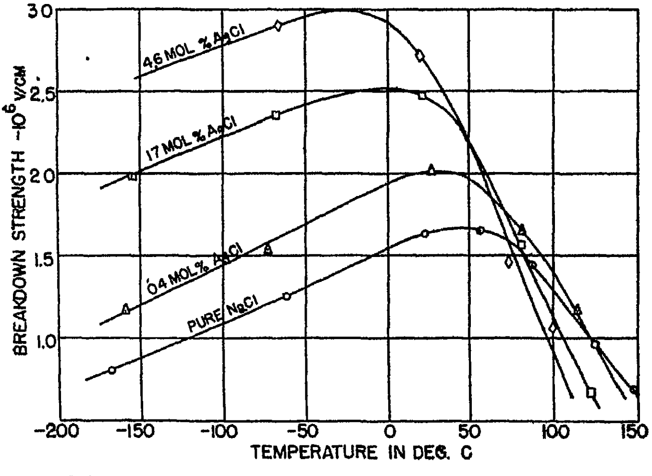


FIG. 2.—Breakdown strength of mixed crystals NaCl–AgCl as a function of composition and temperature.

would seem that for molecules with moments of inertia of 500×10^{-40} g. cm.^{2} a correction of the order of magnitude of 5 % is required at a frequency of 50,000 megacycles/sec. (6 mm. wavelength). It seems desirable to extend current theories of dielectric loss to include inertia terms ; it may be noted that Rocard assumed that Stokes' law was applicable to individual molecules.

Dr. S. Whitehead (*London*) (*communicated*) : Part of Mr. Garton's comment on the paper by Gevers and Du Pré can be established alternatively and rather more rapidly and fundamentally, by using the results of the writer's paper, " A Note on the Analysis of Dielectric Measurements ". Now $\omega(\epsilon' + j\epsilon'')$ is a Kirchhoff function and, if $\epsilon'' \ll \epsilon'$, may be replaced by $\omega\epsilon' e^{j \tan \delta}$, whence from the Hilbert transform relation between tan δ and log $\omega\epsilon'$,

$$\log \omega\epsilon' = -\frac{2}{\pi}\int_0^\infty \frac{\tan \delta}{\omega - u}\, du = -\frac{2 \tan \delta}{\pi}\int_0^{l/\tau_0} \frac{du}{\omega - u}$$

if tan δ varies little with frequency up to $1/\tau_0$ and is thereafter negligible as has been assumed. If (with Mr. Garton) $0 < \omega\tau_0 \ll 1$,

$$\frac{1}{\epsilon,}\frac{\partial\epsilon'}{\partial T} = \frac{2 \tan \delta}{\pi}\,\frac{d \log \tan \delta}{dT} \log \frac{1}{\omega\tau_0},$$

is obtained by differentiation and evaluation where T = abs. temperature. Thus the formula of Gevers and Du Pré,

$$\frac{1}{\epsilon'}\frac{\partial\epsilon'}{\partial T} = \frac{2 \tan \delta}{\pi T} \log \frac{1}{\omega\tau_0}$$

means that tan δ varies little with ω up to $1/\tau_0$ but exponentially with the inverse of the temperature. These features are often encountered experimentally, and are common features of several of the theories in current discussion.

C. BREAKDOWN.

ELECTRIC BREAKDOWN OF SOLID DIELECTRICS.

By A. von Hippel.

Received 20th February, 1946.

I. Experimental Facts.

(A) Breakdown Field-strength and Pre-breakdown Currents.—A thin plate of a solid dielectric, when properly shielded in a sample holder and subjected to homogeneous d.c. fields of slowly increasing strength, goes through several stages of conduction in succession.[2] In weak fields, the

---

[1] von Hippel, *Z. Physik*, 1931, 68, 309 ; 1932, **75**, 145 ; *Ergeb. exact. Naturw.* 1935, **14**, 79 ; *J. Appl. Physics*, 1937, 9, 815.

[2] von Hippel, *Physic. Rev.*, 1938, **54**, 1096.

[3] Rogowski, Flegler and Tamm, *Arch. Elektrot.*, 1927, **18**, 479.

[4] Haworth and Bozorth, *Physic. Rev.*, 1932, **39**, 845 ; *Physics*, 1934, **5**, 15.

[5] von Hippel, *Z. Physik*, 1932, **75**, 145 ; *Z. Electrochem.*, 1933, **39**, 506.

[6] von Hippel, *Z. Physik*, 1934, 88, 358 ; von Hippel and Lee, *Physic. Rev.*, 1941, **59**, 824.

[7] Buehl and von Hippel, *Physic. Rev.*, 1939, **56**, 941.

[8] Austen and Hackett, *Nature*, 1939, **143**, 637 ; Austen and Whitehead, *Proc. Roy. Soc., A*, 1940, **176**, 33.

[9] von Hippel and Maurer, *Physic. Rev.*, 1941, **59**, 820.

current at constant voltage decreases with time; in higher fields, this decay becomes less and less pronounced and may even be replaced by an initial increase of conductivity; finally, at a critical field-strength the current rises abruptly by many orders of magnitude corresponding to the breakdown of the dielectric. This breakdown, from onset to finish, requires only about 10^{-7} seconds.[3] Somewhat below the breakdown field-strength, E_{max}, the current may become unsteady. Single shots can be heard after amplification and can be measured as noise, such fluctuations growing in number and intensity with voltage and diminishing at constant voltage with time.[2, 4]

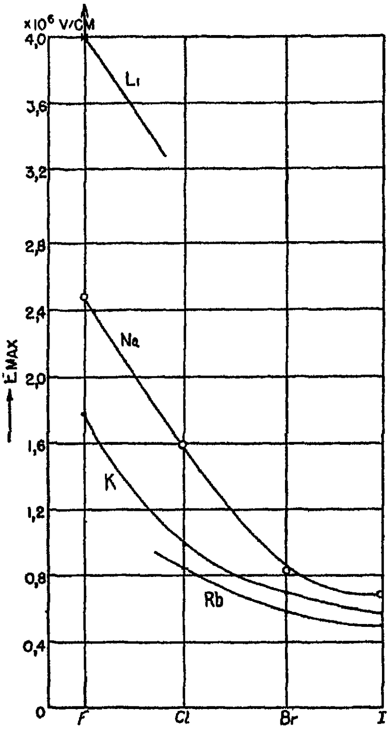


Fig. 1.—Breakdown strength of the alkali halide crystals at room temperature.

The breakdown field-strength of a material, measured under conditions that eliminate edge effect, is a function of structure, composition and temperature. For alkali halides it decreases systematically from Li to Rb and from F to I (Fig. 1),[5] is influenced by addition agents (Fig. 2),[6] and passes through a maximum near room temperature.[6, 7, 8] Passing from the crystalline to the amorphous state, the breakdown strength at low temperature increases and the temperature dependence in this range[9, 11] disappears, as is indicated by data on quartz (Fig. 3).[9] E_{max} is independent of the sample thickness down to 10^{-4} cm. or less, as shown in amor-

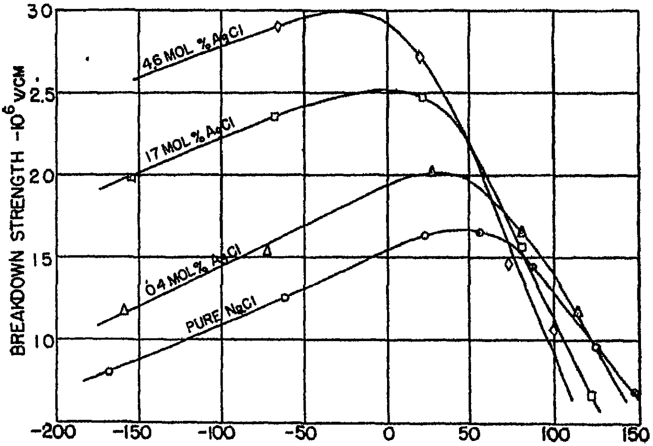


Fig. 2.—Breakdown strength of mixed crystals NaCl–AgCl as a function of composition and temperature.

phous materials,[10, 11] in rocksalt,[5] and in mica.[10, 12] While a slight dependence of $E_{\max}$ on the orientation of the crystal axis in the field can be observed in quartz (Fig. 3),[9] the breakdown strength of NaCl appears to be independent of its orientation.[13]

The average dark current below breakdown rises rapidly with field-strength (Fig. 4).[3, 14] Referred to the same field-strength, the current increases from LiF to RbI, but replotting of the data on a reduced scale, $E/E_{\max}$, makes the characteristics of the different alkali halide crystals practically coincide (Fig. 5).[3] If light can be absorbed in these crystals by "F" centres, a photo-current is observed which reaches saturation as the field-strength increases (Fig. 6).[3] The dark current increases rapidly with temperature, but apparently not in a simple exponential fashion.[7, 11, 12]

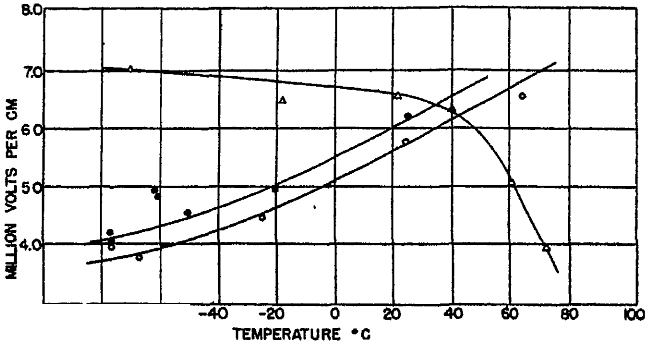


FIG. 3.—The electrical breakdown strength of silica. Quartz, -●- field parallel to axis, -○- field perpendicular to axis; silica glass, -△-.

(B) Breakdown Paths and Their Crystallographic Orientation.—If secondary destruction of the breakdown channel is avoided by limiting the discharge current to very low values, a distinct crystallographic orientation of the breakdown paths in single crystals can be observed.[15, 16, 17] Simple direction laws become discernible if the phenomenon is investigated in homogeneous fields.[16] In NaCl at room temperature, with the field oriented in the [100] direction, the breakdown path initially propagates from the anode towards the cathode in the [110] direction, but is likely to change to [111] paths before reaching the cathode.[16] Thus characteristic tetrahedral pyramids are formed, with their bases resting on the cathode plane (Fig. 7). For the sequence of the alkali halide crystals, the breakdown directions obtained at room temperature are summarised in Table I.[5]

The effect of temperature on these breakdown directions in the alkali halides has been investigated recently by J. W. Davisson.[18] Table II indicates the striking dependency observed.

The results obtained with the Na, K and Rb salts suggest the existence of a sequence: random (R)→[100]→[111]→[110], which is followed by

---

[10] Joffé, *The Physics of Crystals*, McGraw-Hill Book Co. (1928); Alexandrow and Joffé, *Physik. Z. Sowjet.*, 1932, **2**, 527.

[11] Morris Thomas and Griffith, *J. Inst. Elec. Eng.*, 1942, **85**, 487.

[12] Hackett and Morris Thomas, *J. Inst. Elec. Eng.*, 1941, **88**, 295.

[13] von Hippel and Davisson, *Physic. Rev.*, 1940, **57**, 156.

[14] Maurer, *J. Chem. Physics*, 1941, **9**, 579.

[15] Inge and Walther, *Z. Physik*, 1930, **65**, 830.

[16] von Hippel, *Z. Physik*, 1931, **67**, 707.

[17] Lass, *Z. Physik*, 1931, **69**, 313.

[18] Davisson, Ph.D. thesis, *The Orientation of Electrical Breakdown Paths in Single Crystals*, Laboratory for Insulation Research, Massachusetts Institute of Technology, 1943.

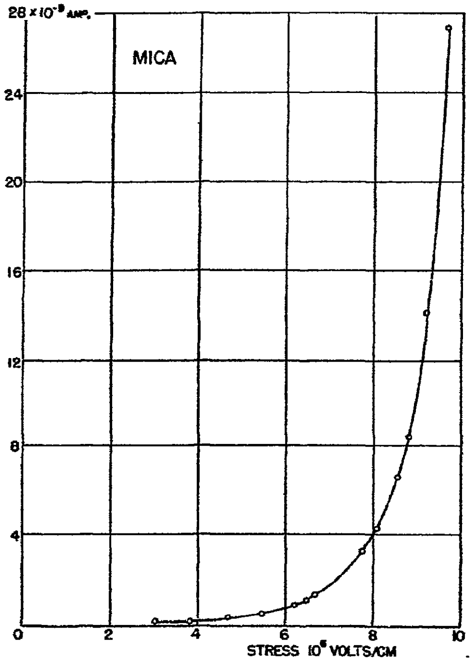


FIG. 4.—Pre-breakdown current in mica.

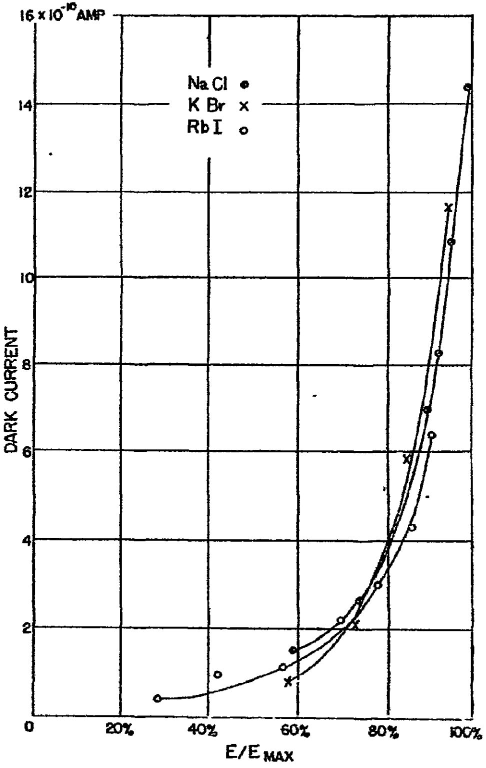


FIG. 5.—Pre-breakdown currents in three alkali halide crystals.

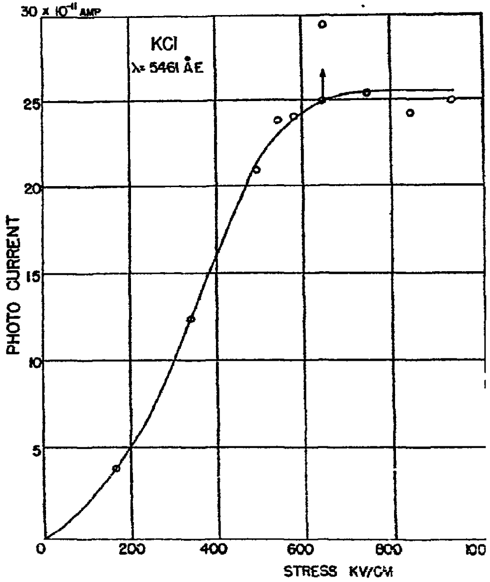


FIG. 6.—Photo-current in KCl discoloured by "F" centres.

the direction of breakdown as lattice energy or temperature increases (Table II). In the Li salts a star pattern (S) may be obtained, formed by 12 non-primitive directions, these directions being [133] for LiF at 150° C., [122] for LiF at 500° C., and [552] for LiCl at room temperature and above (Fig. 8). Increasing lattice energy and temperature apparently tend to remove the degeneracy of the breakdown directions. Thus the least degenerate and most general star pattern [*xyz*], with 24 branches, might be expected in LiF near its melting-point.

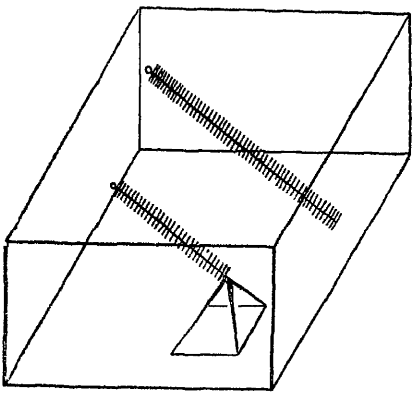


FIG. 7.—Breakdown of NaCl (paths starting at the anode in the [110] direction and splitting near the cathode into [111] paths).

The effect of addition agents upon the breakdown directions is similar to the effect of an increase in temperature.[6, 18] Thus NaCl containing 4·6 mol. per cent. of AgCl, shows [110] paths at liquid air temperature, but

TABLE I.—BREAKDOWN PATHS IN THE ALKALI HALIDES

Crystal.	*Appearance.*	*Direction.**
LiF	sharp	first [110] second [111]
NaF	,,	
NaCl	,,	
KCl	,,	[110] and [100]
RbCl	,,	
NaBr	slightly blurred	[110], [100] and R(andom)
KBr	blurred	[100] and R
KI	,,	
RbBr	,,	
RbI	,,	

* Field in [100] direction ; " first " and " second " refer respectively to the path starting at the anode and to the direction preferred near the cathode.

TABLE II.—BREAKDOWN PATHS IN THE ALKALI HALIDES AS FUNCTION OF TEMPERATURE

	T°C	F	Cl	Br	I
Li	—180	[110]	—	—	—
	20	[110] and S(tar)	[110]	—	—
	200	S	S	—	—
Na	—180	[111]	[100] and R	R	—
	20	[111] and [110]	[110] and [111]	R	—
	200	[110] and [111]	[110]	[110] and [111]	—
K	—180	[100] and [110]	[100]	R	R
	20	[110] and [100]	[100]	[100]	R
	200	[110]	[100] and [110]	[100]	[100] and R
Rb	—180	—	[100] and R	R	R
	20	—	[100] and [110]	[100] and R	R
	200	—	[100] and [110]	[100] and [110]	[100]

reduction of the AgCl content to 1% allows the [100] directions to reappear as companion paths.

These direction laws for breakdown through the volume of a crystal reappear in the directions which discharge paths may take on its surface.[19] Breakdown paths penetrating a NaCl crystal in the [110] direction also form [110] surface forks on the cathode. On the (100) surface of KCl, [100] paths are obtained at room temperature, and on KBr surfaces random orientation prevails. By addition of KBr or RbCl to KCl the surface discharge is compelled to develop in the [110] directions as the rule of sequence would require.

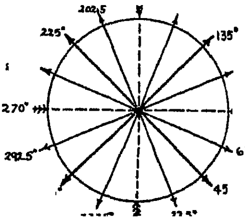


FIG. 8.—Projection on the (100) plane of the 12 paths star pattern in LiCl.

The direction laws discussed above refer to breakdown paths developing from the anode, as they do in homogeneous fields or in non-homogeneous arrangements not favouring the cathode.[20] If the breakdown is forced to begin at the cathode, as by a point-plane arrangement,[17] [100] paths dominate in NaCl at room temperature. The direction laws for cathode paths appear to be related to those for anode paths if account is taken of the fact that for the development of a discharge from the cathode higher field-strength is required. Increased field-strength, as experiments on anode paths indicate, affects the breakdown directions in the same way as does lowering of the temperature.

By investigating a sequence of lattices, ranging from ionic to covalent, it became apparent that the breakdown directions are not governed by the bond type of the crystal but by the laws of macroscopic symmetry.[18] The characteristic pattern types observed in barite,[21] for example, reappear in aragonite as well as in sulphur, indicating that they are a feature of the orthorhombic system.

II. Theoretical Discussion.

The speed with which an electric breakdown develops classifies the process as an electronic phenomenon.[16] Its interpretation revolves, therefore, around the question of how electronic conductivity of sufficient magnitude may arise in dielectrics under the influence of high fields. Three possibilities have been presented : (*a*) electrons from the filled band of the insulator leak out into the conduction band in destructive number by the " tunnel effect " when the field-strength reaches a critical value. This has been proposed by Zener [22] and discussed more in detail by Franz.[23] (*b*) Excess electrons, produced by some ionising process or introduced into the structure from the outside, are accelerated in the applied field after overcoming the friction barrier of lattice vibrations, and produce impact ionisation, avalanche formation and breakdown. This process has been suggested by the author [1] and has been mathematically treated by Seeger and Teller [24] and by Franz.[23] (*c*) Surplus electrons, as in (*b*), initiate the breakdown. However, even in the absence of an applied field these electrons are not stopped by the vibration barrier, but are thermally excited to energies capable of producing impact ionisation. The field upsets the balance between ionisation and recombination rates, and increases the number of conducting electrons, without avalanche formation, until breakdown results. This process is the basis of the theory of Froehlich.[25]

---

[19] von Hippel, *Z. Physik*, 1932, **75**, 169 ; 1934, **88**, 358.
[20] von Hippel, *ibid.*, 1936, **98**, 580.
[21] von Hippel, *ibid.*, 1931, **68**, 309.
[22] Zener, *Proc. Roy. Soc.*, *A*, 1934, **145**, 523.
[23] Franz, *Z. Physik*, 1939, **113**, 607.
[24] Seeger and Teller, *Physic. Rev.*, 1938, **54**, 515 ; 1939, **56**, 349.

Zener's calculations presuppose that the electrons in the filled band of the insulator can be exchanged freely through the material. However, in the case of the alkali halides at least, the electron clouds of the halogen ions apparently do not overlap sufficiently to form a continuous band. Furthermore, the breakdown of amorphous materials does not seem to differ in principle from that of crystals; hence the periodicity of the structure, on which the band picture relies, is not a prerequisite of the actual mechanism.

The following discussion, which elaborates the ideas of the author, shows that the experimental facts may be readily understood on the basis of (*b*) and that Froehlich's theory appears to be untenable.

(A) The Breakdown Condition.—Fig. 9 indicates schematically the probability with which electrons of various kinetic energies excite vibrations in a dielectric material. Each value of kinetic energy can be coordinated with a certain field-strength, E, for which the net loss suffered by the excitation of vibrations balances with the gain resulting from the electric field. Any field-strength below the maximum of the vibration barrier is associated with two kinetic energy values, one for slow electrons corresponding to stable current conditions, the other for fast electrons designating a velocity above which an electron will be accelerated in the applied field until stopped by inelastic impact leading to electronic excitation and ionisation. Our breakdown criterion is that E_{max} represents the critical field-strength at which slow electrons cross the barrier with sufficient probability to generate electron avalanches by impact ionisation. E_{max} in Fig. 9 is associated with the top of the barrier but may actually lie somewhat lower.

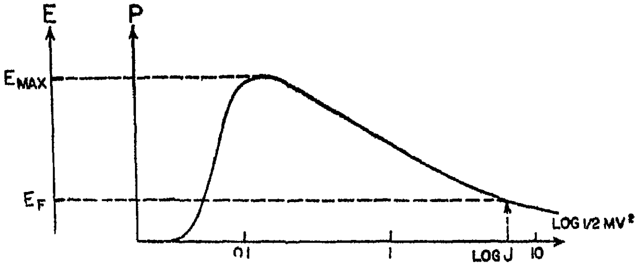


FIG. 9.—Schematic representation of the probable energy loss per unit distance, P, and of the equilibrium field, E, as a function of the kinetic energy of excess electrons.

The maximum energy loss to the friction barrier is that due to the excitation of the vibration modes of maximum frequency. In the alkali halide crystals these modes are closely related to the optical *Reststrahl* frequency, and the product of E_{max} and lattice spacing should therefore to a first approximation increase proportionately to the *Reststrahl* frequency.[1, 26] This has been confirmed by our measurements.[1]

According to Froehlich,[25] in the absence of an external field, excess electrons exist in the conduction bands of the dielectric in an approximate Maxwell distribution with energies between zero and the ionisation energy J of the crystal. This distribution, if disturbed by the removal of fast electrons, is re-established in about 10^{-9} seconds. Froehlich defines as breakdown field-strength, E_F, the field which allows electrons of the energy $1/2\ mv^2 = J$ to be accelerated (Fig. 9). It is, however, the opinion of the author that in the absence of an applied field the friction barrier may be considered to be opaque to electrons which can couple sufficiently strongly with the lattice, as is certainly the case for the materials under discussion. Thus the fast electrons on which Froehlich bases his breakdown criterion do not normally exist. This question will be further considered in the following section.

---

[25] Froehlich, *Proc. Roy. Soc. A*, 1937, 160, 230; 1939, 172, 94; 1941, 178, 493; *Physic. Rev.*, 1939, 56, 349.

[26] Froehlich and Mott, *Proc. Roy. Soc.*, *A*, 1939, 171, 496; Seeger and Teller, *Physic. Rev.*, 1939, 56, 352.

(B) Slow Electrons in Dielectrics.—Conduction bands in dielectric crystals are well defined while empty, but electrons occupying them tend to destroy by polarisation the periodicity required for their existence. This polarisation is a function of the surroundings of the electron and of its velocity and can be expressed in terms of an " effective " dielectric constant. As the velocity of the electron increases, the polarisation which it induces decreases because of the inertia of the nuclei; and the band picture becomes more rigorous. If an excess electron is surrounded by many others as in highly conducting semi-conductors or in metals, these electrons screen the lattice from the polarising action of the electron under consideration, thus maintaining the band structure.

In dielectrics, however, the situation normally prevails that there are insufficient conducting electrons for effective screening, and that the electrons are slow enough to permit nuclei to change their positions while the space charge field acts. Consequently there develops around the electron a distorted lattice area, a potential trough, consisting of induced dipoles in non-polar materials and deepened in polar materials by the orientation of permanent dipoles and by the displacement of ions. The situation is somewhat analogous to the atmosphere surrounding ions in solution, but with the essential difference that in our case the particles forming the atmosphere are locally bound and cannot follow the motion of the electron adiabatically. Consequently, as soon as such a distorted area begins to form, it slows the electron down by dissipating some of its energy in lattice vibrations. The slower electron in turn develops a denser atmosphere increasing the friction, and in short order the electron will find itself trapped in a potential trough of a depth given by the deviation of the effective static from the effective optical dielectric constant.

This trapping mechanism visualised by the author,[1, 27] does not violate quantum mechanics. Landau[28] has suggested that slow electrons may not be trapped on account of their long wavelength and that first a speed-up process is required. However, the effective mass of the electron is increased by the polarisation of its surroundings causing the wavelength to shrink and allowing localisation as the electron slows down.

Consequently we normally do not expect to find electrons moving freely in dielectrics with a Maxwell distribution of velocities. Their motion will be an intermittent one, alternating between times of movement, t_m, and times of rest, t_r, and the length of the rest period will be given by the thermal activation energy, U, of the trapped state,

$$t_r = a\,e^{+\frac{kT}{U}} \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

The drift velocity, v, in a weak electric field will be represented by

$$v = \frac{t_m + t_r}{\bar{w}}, \qquad . \quad . \quad . \quad . \quad . \quad (2)$$

where $\bar{w}$ is the mean displacement distance between two trapping acts. If $t_r \gg t_m$, the drift velocity will decrease exponentially with the reciprocal of the temperature

$$v = \frac{\bar{w}}{a}\,e^{-\frac{kT}{U}} \qquad . \quad . \quad . \quad . \quad . \quad (3)$$

---

[27] von Hippel, *Z. Physik*, 1936, **101**, 680; *J. Chem. Physics*, 1940, **8**, 605.
[28] Landau, *Physik. Z. Sowjet.*, 1933, **3**, 664.

This has been observed by Pohl and his co-workers [29] for excess electrons in alkali halide crystals by measuring the motion of the " F " centres.[30] It has also been found, by photo-electrically releasing N_0 electrons at lower temperatures at various distances, x, from the anode,[31] that the trapping occurs with statistical probability, *i.e.*,

$$N = N_0 e^{-\frac{\bar{w}}{x}}, \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and that the mean trapping distance increases proportionately to the field-strength, E, applied,

$$\bar{w} = cE. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

For the alkali halides, $\bar{w}$ is about 10^{-7} cm. in a field of 1 volt/cm.[32] This explains the high field-strength required in our experiments for saturating the photo-current (Fig. 6).

From (5) it follows that the expressions (2) and (3) for the drift velocity will no longer be valid when $\bar{w}$ becomes larger than the distance, d, between cathode and anode. In this region the drift motion will cease to be intermittent and will become independent of t_r. By measuring the current density

$$j = Nev. \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

and its dependence on field-strength, temperature, structure, and composition we hope to learn more about the trapping mechanism at low field-strengths and about the modes of motion of the electrons as the field increases. It may be possible to explore the vibration spectrum of the dielectric by breaks in the current-voltage characteristic. Experimental difficulties arise as a result of field emission from the cathode, space charge effects, and the onset of avalanche formation. We are now engaged in this study.

(C) Directional Breakdown as an Interference Phenomenon.—Breakdown according to our picture will be initiated by slow electrons having energies not far surpassing the *Reststrahl* quantum. The wavelength of an electron of 10^{-1} e. volt is about 4×10^{-7} cm. long and covers about 7 periodicity distances in NaCl. Thus only high order interference could be produced. However, the possibility of setting up an interference pattern of such high order is destroyed by the electron itself through its space charge action, which distorts the lattice and obliterates the long range periodicity. It is thus understandable that the breakdown strength of NaCl is not measurably influenced by the orientation of the crystal relative to the field.[13] However, a slight influence of the axis orientation can be measured in quartz (Fig. 3); [9] because a much higher breakdown strength and consequently higher speeds of the electrons are involved, and also because of the pronounced anisotropy of the quartz crystal.

When the excess electrons pass the top of the vibration barrier in the breakdown field and become accelerated, their wavelength shrinks, and the lattice distortion diminishes. Hence it should be expected that the wave-fields of these fast electrons set up interference patterns scattering the electrons in preferential lattice directions as determined by the macroscopic crystal symmetry. This we have observed.

Superposed on this " Laue " interference one should expect an interference phenomenon corresponding to the " Ramsauer " effect, a resonance scattering of the electron waves by the electron clouds of the individual lattice ions. A " fine-structure analysis " as for X-rays is indicated, which should for instance explain why the [111] direction appears only in the sodium halides (Table II).

---

[29] Pohl, *Proc. Physic. Soc.*, 1937, 49, 3.

[30] Stasiw, *Nachr. Ges. Wiss. Göttingen*, 1932, 26; 1933, 50; Smakula, *ibid.*, 1934, 4.

[31] Flechsig, *Physik. Z.*, 1931, 32, 843; Hecht, *Z. Physik*, 1932, 77, 235.

[32] Glaser, *Nachr. Ges. Wiss. Göttingen*, 1937, 2.

A quantitative discussion of the direction laws is at present being attempted on the assumption that the breakdown paths mark the direction in which the electrons travel for the major distance between two ionising impacts. Of the paths allowed by the crystal symmetry, the one which will actually appear depends on the acceleration with which the electrons traverse the different velocity regions, and therefore on the shape of the friction barrier between its maximum and the ionisation energy (Fig. 9). The probability of exciting vibrations will increase with the coupling strength between electron and lattice and with temperature and addition agents disturbing the periodicity of the lattice. These factors increase the distance between two ionising impacts and permit breakdown directions to appear corresponding to higher electron velocities. Increase of the ionisation energy should act in the same sense while increase of the field-strength should influence the selection of patterns in the opposite direction, as observed. Random paths will form, if the acceleration is too fast for the build-up of an effective interference field.

How the paths actually develop through the action of the positive space charge left by the avalanches has been discussed elsewhere.[1] It appears that we can make visible and can study in detail the modes of conduction in solid dielectrics through an investigation of the breakdown of crystals.

Summary.

The electric breakdown of solid dielectrics, as the author has suggested a number of years ago, may be due to electrons which are accelerated after overcoming the friction barrier interposed by the vibrations of the material. The experimental facts since accumulated permit a more detailed discussion of this process. It appears that rather complete information may be obtained from breakdown experiments as to the probability with which electrons of various velocities excite lattice vibrations. The energy loss suffered by slow electrons, which are not able to cross the top of the friction barrier, may be derived from the current-voltage characteristic preceding breakdown. The breakdown field-strength itself defines the maximum of the excitation function. The crystallographic orientation of the breakdown paths is determined by faster electrons which have traversed this maximum. These electrons create characteristic interference patterns by their wave field while being accelerated towards ionisation, and the breakdown directions mark paths of lowest friction leading through this interference field. Experiments and calculations, interrupted by the war, are being resumed to check further the validity of this interpretation.

Laboratory for Insulation Research,
Massachusetts Institute of Technology,
Cambridge, Massachusetts.

GENERAL DISCUSSION.

Dr. H. Fröhlich (*Bristol*) said: (i) Von Hippel's consideration II(B) no doubt gives a picture of the average motion of an electron. This, however, can not be used as an objection against the existence of a Boltzmann law for the energy distribution of electrons. Even so, the number of fast electrons will be extremely small in the absence of an external field. This is different, however, in the presence of a field F. I have calculated the energy distribution for this case and found that the deviations from the Maxwell distribution becomes appreciable when $eFl < 1$ (l is the mean free path of electrons which depends on the electronic energy). Details and further results will be published elsewhere.*

---

* *Proc. Roy. Soc., A,* 1947.

(ii) I should like to mention that the experimental results on the dependence of breakdown strength on temperature and on the concentration of foreign atoms can easily be understood in terms of the mean free path of fast electrons. No explanation of these important effects can be offered if the breakdown strength is determined by the behaviour of slow electrons.

Dr. R. W. Lunt (*Birmingham*) said : I should like to be allowed to change the spirit of the discussion from the refinements of theory to the equally important consideration of how far existing theory serves the practical and very important problems of insulation in the laboratory and in industry.

It will have been noticed that all of von Hippel's communication relates to certain simple crystals about the structure of which a considerable amount is known. What is so greatly to be desired is the extension of our knowledge both theoretically and experimentally to what may be called " classical " insulators such as sulphur and diamond and also to the insulators of engineering importance, for example, paraffin wax, transformer oil, oil-impregnated paper and polythene. My impression is that the same principles apply in the theoretical treatment but that, because the wave-mechanical analysis is more tedious and complicated in the cases of diamond and sulphur than, for example, in that of sodium bromide, no theoretical physicist has yet thought that the insulators commonly used are important enough to merit attention.

In the case of oil-impregnated paper there is considerable experimental evidence indicating that the incidence of failure of the insulation in high fields of the order of 10^5 v. cm.$^{-1}$ increases markedly as the mean temperature of the material is increased in the range 20 to 100° C. ; and that in the initial phase of failure the system behaves as a generator of oscillations extending over a wide range of high frequencies. The variation of electric strength of the insulation with temperature is thus different in sense from that found for simple ionic crystals such as sodium bromide.

It is therefore of interest to enquire whether such simple ionic crystals as well as the insulators used in engineering possess characteristic absorption frequencies in the frequency range about 10^6 to 10^9 c./sec. such as have been reported for some other materials. For in this case it appears that the electric strength would vary with the frequency of the applied field and pass through a minimum in the neighbourhood of an absorption maximum ; and furthermore there is also the possibility that the oscillations generated in an insulator in the initial phase of failure may be preferentially absorbed and that this absorption offers a partial explanation of the progressive development of failure.

It is pertinent in any discussion of ultimate electrical strength to raise the nice question as to how far the experimental observations are confused by the incidence of cosmic showers. My recollection is that the first maximum of probability for the interception of cosmic particles and their conversion into cosmic showers, occurs for a stopping power roughly equivalent to that of 0·5 cm. of lead. Since usual engineering practice in the design of both high voltage insulated cables and high voltage transformers is to surround the insulation with metal of such a thickness that the stopping power approximates to that of 0·5 cm. of lead, it is obvious that the standard practice in design provides for a relatively high generation of cosmic showers in the insulation. Furthermore, because the stopping power of the insulation, for example, oily paper or transformer oil, is roughly three orders of magnitude greater than that of air (in which gas cosmic showers have usually been studied), the range of the shower tends to be confined to the insulation ; in consequence the concentration of fast electrons forming the shower is much more intense than that associated with a cosmic shower in free air so that on this account also the engineering design appears to favour the conditions for breakdown. And in the experimental determination of the electric strength of such materials as sodium bromide, reference has just been made by the previous speaker to the use of massive electrodes: these must also favour the occurrence of showers in the material under examination.

Since insulating material in engineering applications is inevitably and uncontrollably liable to be traversed by cosmic particles and mesotrons, and since the containing metal and the stopping power of the insulating material jointly contribute to the development of dense showers in the insulating material, it is possible that the incidence of such showers sets an upper limit to the effective electric strength for engineering practice. It is therefore a matter of prime importance to determine the extent to which such showers prescribe the electric field strength to which any given insulating material can be safely and continuously subjected, especially when a metal tank or sheath is provided to contain the insulating material. For if the effect of showers is considerable, there would be a limit beyond which it would be uneconomic to attempt to increase the electric strength *in the absence of showers*, for example, by diminishing the chemical impurities in the material.

Dr. S. Whitehead (*London*) (*communicated*): Dr. Lunt will find in the experimental studies given in a series of papers by the present writer and his colleagues, Dr. Austen, Mr. Thomas, Dr. Hackett and Dr. Pelzer that the theory given by Fröhlich gives reasonably good concordance for mica and quartz, paraffin wax and polythene. Supplementary hypotheses are required for the vinyl co-polymers and more complex insulating varnishes, but the relative importance of ordered structure and molecular complexity has not been established. However, the variation with temperature mentioned by Dr. Lunt has little to do with electronic breakdown since he refers to fields far below the intrinsic electric strength where other mechanisms enter. The other mechanisms so far established are thermal breakdown, internal discharges and a species of electro-chemical deterioration. The first and last must nearly always lead to a decrease of electric strength with temperature. Internal discharges can be associated with such a variation but not as a direct consequence. These discharges are, however, normally the cause of the oscillations to which Dr. Lunt refers; since they provoke the natural oscillations of the observing circuit. It has been shown that dielectric breakdown at 10^6-10^9 c/s. is mainly thermal, and varies with temperature and frequency as postulated by theory except where internal discharges occur and in many circumstances such discharges cause thermal breakdown by the increased energy dissipation which they cause either directly or indirectly as by carbonisation.

The theory of electronic breakdown requires no more initiating electrons than a gaseous discharge and the energy limitation is only a few electron-volts as compared with the 20 e.v. or more in atmospheric discharges. The latter may therefore be regarded as more susceptible to the influence of cosmic rays. This influence has been claimed to have been observed but only when very severe conditions are applied to remove other effects which normally mark the former influence. It may, therefore, be concluded that it would be all the more difficult to find conditions in which cosmic particles materially influenced dielectric breakdown.

Dr. A. E. W. Austen (*London*) said: At moderate stresses and ordinary temperatures the diminishing current flowing in most, if not all, good dielectrics for many seconds and sometimes many hours after the sudden application of a d.c. voltage, is completely reversible. It is merely another manifestation of the same class of phenomena as cause a finite power loss on sinusoidal voltages and a steady increase in dielectric constant with decrease of frequency, and which are discussed in most of the other papers in this series. It is important to distinguish between such reversible currents and true conduction currents, since the former cannot be due to the permanent, as compared with elastic, displacement of ions.

On suddenly applied direct voltage, the conduction current may become appreciable compared with the reversible current after a considerable time.

Constancy of current with respect to time is a misleading criterion of conduction current and it may best be established by comparing charging current with discharging current. At the times which, as far as can be discerned from his published accounts, Von Hippel is concerned with, the conduction current at moderate stress would be small compared with the reversible current. Von Hippel, however, worked at high stress. Now it has been established [1] that the reversible currents do not show any appreciable non-linearity (*i.e.* current/voltage remains constant) up to stresses of order 1/5 the electric strength. Thus if Von Hippel's currents are conduction currents it requires that the conduction current, which is small compared with the reversible current at 1/5 $E_{\max}$, at higher stresses increases more rapidly than the stress without there being any corresponding abnormal increase in the reversible current. In my opinion this is a phenomenon sufficiently surprising to require more detailed experimental evidence than Von Hippel has so far given.

Tests at high electric stress are best made with only the test dielectric and the electrode metal present in regions of high stress. If a third substance, *e.g.* an immersion medium, is present, discharges in it or at interfaces are very difficult to avoid. Electrical " noise " is a characteristic indication of such discharges and a decrease of frequency of occurrence with time such as is mentioned by Von Hippel is also a familiar feature of edge discharges on d.c.

Dr. Austen (*London*) (*communicated*): Dr. Lunt's suggestion that oil-impregnated paper at high stress " behaves as a generator of oscillations extending over a wide range of high frequencies " would appear to derive from some early American work,[2] though the idea has reappeared not infrequently since. It is probably a main cause of the air of mystery that tends to surround the subject of internal discharges. Dr. Hackett and I have shown [3] that the phenomenon consists of the sudden discharge of a local region of the dielectric of lower electric strength than the dielectric in series with it, which is detectable at the terminals of the test capacitor only as a sudden apparent loss of charge. But if the test capacitor forms part of a resonant circuit, and for the purpose of detection of internal discharges it is often so connected, there will follow a damped oscillation at the resonant frequency. The dielectric is a " generator of oscillations " only in the sense that the voltage form produced, a Heaviside unit function, may for convenience of analysis be said to consist of a series of harmonic components.

---

[1] Austen, *J.I.E.E.*, 1945, 92, (1), 373, 1945.

[2] Tykociner, Brown and Paine, *Univ. of Illinois Eng. Exp. St. Bull.*, 1933, No. 259 and No. 260.

[3] Austen and Hackett, *J.I.E.E.*, 1944, 88, 88 (1).

D. MEASUREMENT TECHNIQUE AND TERMINOLOGY.

THE REPRESENTATION OF DIELECTRIC PROPERTIES AND THE PRINCIPLES UNDERLYING THEIR MEASUREMENT AT CENTIMETRE WAVELENGTHS.

By Willis Jackson.

Received 6th March, 1946.

1. Correlation of the Parameters used in Dielectric Measurements.

Depending on the purpose for which a knowledge of the electrical properties of dielectric media is required and on the method by which they are measured, it is convenient and usual to represent them in different ways. The relationships between the various paramenters involved are not perhaps generally known, and a short statement on the matter may therefore be found useful.

In the first place a dielectric medium may be specified in terms of a dielectric constant (or permittivity) ϵ, a permeability μ, and an effective conductivity σ. During recent years it has become the practice among physicists and electrical engineers to express these quantities in the rationalised M.K.S. system of units, in which system they are given in farads, henrys and mhos per metre respectively, the permittivity and permeability of free-space (vacuum) having the values

$$\epsilon_v = 1/(36\pi) \times 10^9 \text{ farads/metre}$$

and $$\mu_v = 4\pi \times 10^{-7} \text{ henrys/metre.}$$

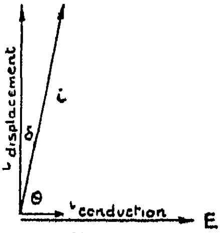


FIG. 1.

Taken *relative* to free-space the permittivity and permeability values of a non-vacuous medium, ϵ/ϵ_v and μ/μ_v are the same as their absolute values on the C.G.S. system.

The current density in amperes/sq. metre associated with an applied electric field strength $e = Ee^{j\omega t}$ volts/metre is then given by

$$\begin{aligned} i &= \sigma e + \epsilon \, \partial e/\partial t \\ &= (\sigma + j\omega\epsilon)Ee^{j\omega t} \qquad . \quad . \quad (1) \end{aligned}$$

$\omega = 2\pi f$ being expressed in radians/sec. The first term denotes a conduction current density in time phase with the applied field, and the second a displacement current density in phase quadrature, as shown in Fig. 1. The resultant current density leads E by θ and $\cos \theta$ is called the *power factor* of the dielectric. The complement of the *phase angle* θ, the angle δ, is known as the *loss angle* and clearly

$$\tan \delta = \sigma/\omega\epsilon \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and is indistinguishable from the power factor for small values of δ.

Equation (1) may be rewritten

$$\begin{aligned} i &= j\omega(\epsilon - j\,\sigma/\omega)Ee^{j\omega t} \\ &= j\omega(\epsilon' - j\epsilon'')Ee^{j\omega t} \quad . \quad . \quad . \quad . \quad (3) \end{aligned}$$

in which $\epsilon' - j\epsilon'' = \bar{\epsilon}$ say, denotes an equivalent complex dielectric constant of the conducting medium. Evidently

$$\sigma = \omega\epsilon'' \text{ and } \tan\delta = \epsilon''/\epsilon'. \quad . \quad . \quad . \quad (4)$$

Alternatively if E is the electric field component of a plane electromagnetic wave travelling through the medium in the direction x, the application of Maxwell's equations shows it to vary with x in the manner $E = E_0 e^{-\gamma x}$ where the *propagation constant* γ is equal to $j\omega\sqrt{\mu\bar{\epsilon}}$. γ is obviously complex for complex $\bar{\epsilon}$ and is usually written $\gamma = \alpha + j\beta$, where α is known as the *attenuation coefficient* and β as the *phase constant.* Their respective units are nepers/metre [*] and radians/metre. On substituting for $\bar{\epsilon}$ in γ, it follows that

$$\frac{\beta^2 - \alpha^2}{\mu\omega^2} \text{ and } \epsilon'' = \frac{2\alpha\beta}{\mu\omega^2} \quad . \quad . \quad . \quad (5)$$

$$\alpha = \omega\left\{\frac{\mu\epsilon'}{2}\left[\sqrt{1 + (\epsilon''/\epsilon')^2} - 1\right]\right\}^{\frac{1}{2}} = \omega\left\{\frac{\mu\epsilon'}{2}\left[\sqrt{1 + \tan^2\delta} - 1\right]\right\}^{\frac{1}{2}}$$

$$\beta = \omega\left\{\frac{\mu\epsilon'}{2}\left[\sqrt{1 + \tan^2\delta} + 1\right]\right\}^{\frac{1}{2}}. \quad . \quad . \quad . \quad (6)$$

Still further

$$\sigma = \frac{2\alpha\beta}{\mu\omega} \text{ and } \tan\delta = \frac{2\alpha\beta}{\beta^2 - \alpha^2} \quad . \quad . \quad . \quad (7)$$

in which it will be remembered $\mu = 4\pi \times 10^{-7}$ for non-magnetic media.

The behaviour may also be represented in terms of a *complex index of refraction* written $\bar{n} = n(1 - \kappa)$ where n is the *refractive index* and κ the *index of absorption.* This is equal to the square root of the complex permittivity, $\bar{\epsilon} = \epsilon' - j\epsilon''$ expressed in C.G.S. units, and, on the M.K.S. system, to the square root of the complex *relative* permittivity, $\bar{\epsilon}/\epsilon_v$. It follows that

$$\epsilon'/\epsilon_v = n^2(1 - \kappa^2)\,; \quad \epsilon''/\epsilon_v = 2n^2\kappa \quad . \quad . \quad . \quad (8)$$

and from (4) that

$$\sigma = 2n^2\kappa \,.\, \omega\epsilon_v = \frac{1}{9 \times 10^9} \,.\, fn^2\kappa \text{ mho/metre}$$

$$\tan\delta = 2\kappa/1 - \kappa^2. \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Further, on equating the alternative expressions for γ, namely, $\alpha + j\beta$ and $j\omega\sqrt{\mu\bar{\epsilon}} = j\omega\sqrt{\mu\epsilon_v} \,.\, \sqrt{\bar{\epsilon}/\epsilon_v}$ and replacing $\sqrt{\bar{\epsilon}/\epsilon_v}$ by $\bar{n}$, it follows that

$$n = \frac{\beta}{\omega\sqrt{\mu\epsilon_v}}\,; \quad n\kappa = \frac{\alpha}{\omega\sqrt{\mu\epsilon_v}} \quad . \quad . \quad . \quad (10)$$

in which, for non-magnetic media, $1/\sqrt{\mu\epsilon_v} = 3 \times 10^8$ metres/sec. Alternative forms of expression for n and $n\kappa$ result on substitution for α and β from (6). It will be appreciated without comment that $n = \lambda_v/\lambda$ where λ_v and λ are the wavelengths in free space and in the medium in question corresponding to angular frequency ω.

2. Resonance Methods of Measurement.

The most convenient parameters in which to express the results obtainable from resonance methods of dielectric measurements are the relative permittivity and tan δ. The principle underlying these methods is as follows. Fig. 2 represents a circuit consisting of a coil of inductance L henrys and resistance r ohms connected in series with a condenser of capacitance C farads, the whole dielectric space of which is supposed filled with material of permittivity ϵ farads/metre and conductivity σ mho/metre at the resonant angular frequency $\omega_r = 2\pi f_r = 1/\sqrt{LC}$.

---

[*] Note that 1 neper = 8·686 decibels.

If a parallel plate condenser of plate area A sq. metres and plate separation d metre be assumed, its capacitance[*] is $C = \epsilon . A/d$ farads. On replacing the material in question by air (strictly by a vacuum) and changing the plate separation to d_v so as to restore the capacitance value, ϵ/ϵ_v is given by d/d_v.

The tan δ value is obtained from the change of Q factor of the circuit which results on making this replacement. The Q factor of a resonant system is defined in its most general form by the relation

$$Q = \omega_r . \frac{\text{energy stored in the system}}{\text{power loss in the system}} \quad . \quad . \quad (11)$$

and it follows that if Q and Q_v are the values of this factor with the lossy test dielectric and air dielectric respectively

$$\frac{1}{Q} - \frac{1}{Q_v} = \frac{1}{\omega_r} . \frac{\text{power loss in the test dielectric}}{\text{energy stored in the system}} \quad . \quad . \quad (12)$$

The stored energy oscillates between the coil and the condenser, and is expressed in terms of the peak value E_m of the electric field strength in the condenser dielectric by $\frac{1}{2}\epsilon(Ad)E_m^2$. The corresponding power loss in the dielectric is given by $\frac{1}{2}\sigma(Ad)E_m^2$, so that (12) reduces to

$$\frac{1}{Q} - \frac{1}{Q_v} = \frac{\sigma}{\omega_r \epsilon} = \tan \delta \quad . \quad . \quad . \quad . \quad (13)$$

from equation (2).

The Q factor of a resonant circuit is most conveniently determined from the shape of its resonance curve. This may be delineated by incremental frequency change about the resonant frequency, in which case Q is obtained as $f_r/\Delta f$, where Δf is the width of the resonance curve at half height, the response of the detector system being supposed square law. If, on the other hand, the resonance curve is delineated at fixed frequency by incremental capacitance change about the resonant value C_r, Q is given by the analogous expression $2C_r/\Delta C$. This is the method adopted by Hartshorn and Ward[1] in the dielectric test set which they developed for measurements over the frequency range 10^4-10^8 cps.; the incremental capacitance change is obtained on a calibrated vernier condenser connected in parallel with the variable parallel-plate test condenser. The basic Q value of the resonant circuit of this test set at its high frequency end is of the order of 500, so that the insertion of a specimen[†] of tan δ value 0·0005, say, produces a 20 % reduction to 400.

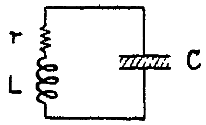


FIG. 2.

The use of "lumped" resonant circuits becomes increasingly impracticable at frequencies above 10^8 cps. and they are then replaced by resonant elements of either transmission lines or waveguides, oscillating in one or other of the many possible characteristic modes.

Transmission lines are seldom operated deliberately in any other than the principal one in which both the electric and magnetic field components lie purely transverse to the line direction. This is illustrated in Fig. 3(*a*) for a half-wave closed coaxial-line resonator, the electric field in which is purely radial and the magnetic field circumferential. For complete air-filling (strictly a vacuum) resonance occurs for lengths equal to any integral multiple n of one-half the energising wavelength $\lambda_0 (= \frac{3 \times 10^{10}}{f}$ cm., where f is in cycles per second). This form of resonator has been used by the writer and his colleagues for measurements at wavelengths of the order of 30 cm. (frequency 10^9 cps.).

---

[1] *J.I.E.E.*, 1936, **79**, 597.

[*] For simplicity edge effects are ignored.

[†] The condenser plate diameter is 5 cm. and a convenient specimen thickness 0·1 cm.

In waveguides, however, a purely transverse mode does not exist, and the operative modes, known as *E* and *H* modes, are characterised by the presence of a longitudinal component of electric and magnetic field respectively. For measurements at wavelengths around 10 cm. (frequency 3×10^9 cps.) use has been made of a resonant waveguide cavity operating in the E_{010} mode.[3] The electromagnetic field distribution characterising this mode is illustrated in Fig. 3*b*; the electric field is purely longitudinal and the magnetic field purely circumferential. In this system

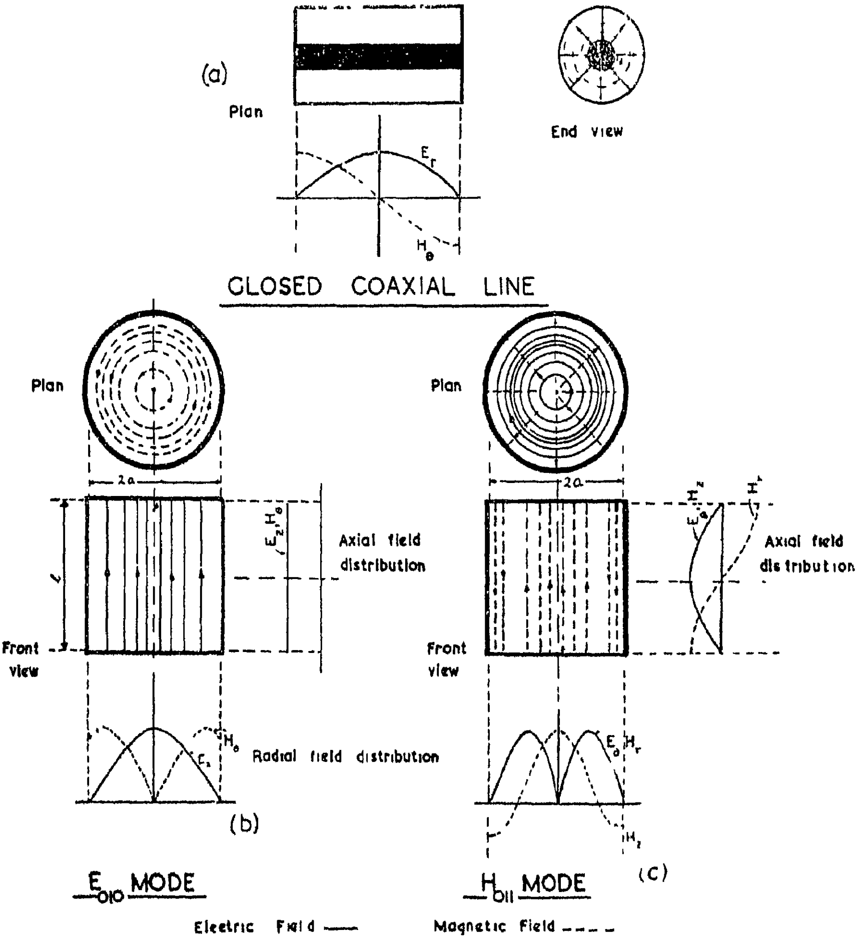


FIG. 3.—Resonator field configurations.

the axial length does not enter into the resonant condition; this is governed entirely by the radius *a*, the resonant energising wavelength for a completely air-filled cavity being $\lambda_0 = 2{\cdot}6125a$.

Fig. 3(*c*) illustrates the field distribution in a resonant cavity operating in the H_{01n} mode, for *n* equal to unity. This mode is characterised by a circumferential electric and by both radial and longitudinal magnetic field components. For an air-filled cavity of radius *a*, the H_{01n} mode can exist only for energising wavelengths λ_0 shorter than a critical, or cut-off, value $\lambda_c = 1{\cdot}64a$. Provided, however, that it can be excited, the dimension which governs the resonant adjustment for this mode is the axial length, and resonance occurs when the length is equal to $n \,.\, \frac{\lambda_g}{2}$, where λ_g denotes the wavelength in the cavity and is related to λ_0 and λ_c by $1/\lambda_g^2 = 1/\lambda_0^2 - 1/\lambda_c^2$. Evidently the closer λ_0 approaches λ_c the longer the resonant half-wave, H_{011}, cavity, a consideration which has to be taken into account in deciding upon the cavity radius most appropriate to operation at a specified energising wavelength. This form of resonator has been used for measurements at wavelengths [2] of 3 cm. and below.

---

[3] For a description of the experimental technique see Horner, Taylor, Dunsmuir, Lamb and Jackson, *J.I.E.E.*, 1946, 93, Part III, 53.

Mathematical expressions for the Q factor of the closed coaxial line, the E_{010} and the H_{01n} resonator systems in terms of their dimensions and of the operating wavelength have been derived in the previous paper.[2] These have been used to provide the comparative figures given in the table for air-filled resonators suitable for operation at a wavelength λ_0 of 10 cm. The cross-sectional dimensions of the coaxial line resonator and the length of the E_{010} one have been chosen somewhat arbitrarily, and it should be mentioned that for the same b/a ratio in the former case, Fig. 3*a*, the Q factor increases in proportion to the increase of radii, while in the latter, Fig. 3*b*, it increases with the length in the ratio $l/(l + a)$.

TABLE I.—COMPARATIVE DIMENSIONS AND THEORETICAL Q FACTORS OF TYPICAL AIR-FILLED AND SILVER-PLATED, COAXIAL LINE, E_{010} AND H_{01} RESONATORS FOR λ_0 = 10 CM.

Coaxial line .	Arbitrary inside radius of outer conductor, b *	2·0 cm.
	Optimum radius ratio b/a	3·59
	Radius of centre conductor, a *	0·55 cm.
	Minimum resonant length $\lambda_0/2$ ($n = 1$)	5·0 cm.
	Q factor	6100
E_{010} resonator .	Cavity radius, $a = \lambda_0/2{\cdot}6125$	3·83 cm.
	Arbitrary axial length †	4·5 cm.
	Q factor	17500
H_{01} resonator .	Cut-off radius $\lambda_0/1{\cdot}64$	6·1 cm.
	Arbitrarily chosen radius a ‡	8·0 cm.
	Minimum resonant length $\lambda_g/2$ ($n = 1$)	7·8 cm.
	Q factor	47500

* With these values of a and b, $\pi(a + b) = 8$ cm., which being less than the specified energising wavelength of 10 cm., no E or H modes can be freely propagated in the system.

† This length excludes the possible excitation of modes other than the desired E_{010}.

‡ With this radius other modes of oscillation are possible but their excitation to significant amount may be inhibited by suitable experimental arrangements.

Even with the coaxial line system the Q factor is very much larger than that of the lumped resonant circuit of the Hartshorn-Ward dielectric test set, and it is evident that a substantial change of Q will result from complete filling of these resonators with material of much smaller tan δ value than the already low figure of 0·0005 used previously for illustration. For example, on the basis of equation (13), the Q factor of the E_{010} resonator would be reduced to about one-half its air-filled value by filling with a material of tan δ = 0·00006. For the measurement of tan δ values of the order of 0·0005 therefore complete filling of the resonator with the test dielectric is unnecessary. This is a great practical convenience, particularly of course with respect to solid dielectrics.

Consideration of the field distributions illustrated in Fig. 3 suggests the shapes and locations of specimen shown in Fig. 4. With these dielectric arrangements the relations connecting the dimensions and electrical properties of the test material with the changes of resonant adjustment and of Q factor which result from its insertion are necessarily somewhat complex,[2] but are not unduly difficult to apply. In the case of the coaxial line and H_{01n} systems these changes may be observed either by variation of oscillator frequency, or by change of resonator length at constant frequency in the manner illustrated in Fig. 4. The latter procedure possesses the considerable advantages that the need for accurate calibration of the oscillator for incremental frequency change required by the former method

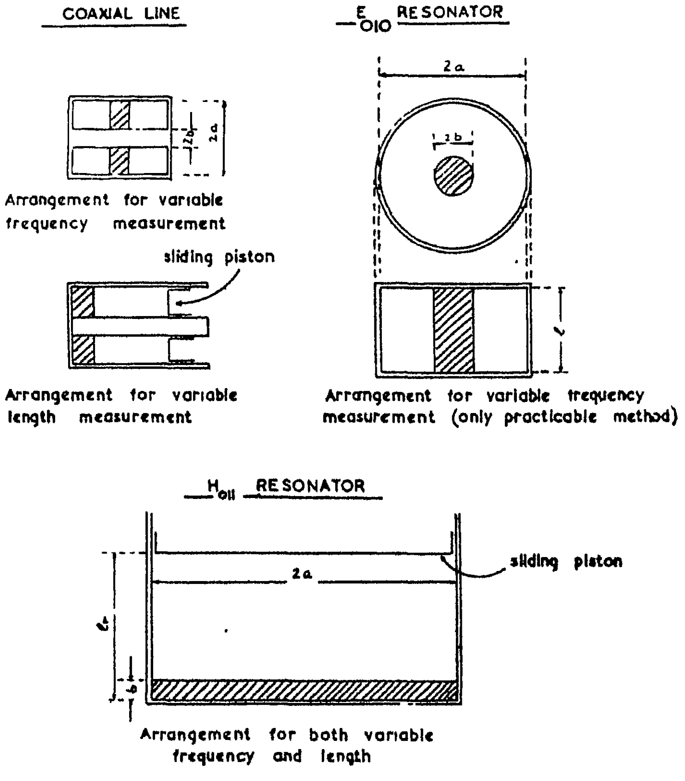

COAXIAL LINE
E_{010} RESONATOR
2a
2b
2b
2a
Arrangement for variable frequency measurement
sliding piston
l
Arrangement for variable length measurement
Arrangement for variable frequency measurement (only practicable method)
H_{011} RESONATOR
sliding piston
2a
l
b
Arrangement for both variable frequency and length

FIG. 4.

does not arise, and the problems associated with the limited frequency range and output variation over their range of centimetre wave oscillators are avoided. With the E_{010} system, however, frequency change is the only practicable procedure, and since this system has a substantially smaller Q factor than the corresponding H_{01n} one, its only significant advantage over the latter lies in its smaller dimensions, a factor which becomes important at wavelengths above about 10 cm. At wavelengths of 30 cm., and above, the excessive dimensions of both the E_{010} and H_{01n} resonators leave no alternative to the coaxial line system.

3. The Standing Wave Method of Measurement.

If a plane electromagnetic wave travelling in a medium of constants μ, $\tilde{\epsilon}_1$, Fig. 5(a), impinges at normal incidence upon the boundary AB between this medium and one of constants μ, $\bar\epsilon_2$, it is well known that wave reflection occurs at the interface and that the magnitude and phase of the reflected wave taken relative to the incident one may be expressed, for the electric field, in terms of a reflection coefficient

$$k = K \, . \, e^{j\psi} = \frac{\sqrt{\bar\epsilon_1/\bar\epsilon_2} - 1}{\sqrt{\bar\epsilon_1/\bar\epsilon_2} + 1} \quad . \quad . \quad . \quad . \quad (14)$$

The interaction of the reflected wave with the incident one leads to a standing wave in medium 1, and from the shape and position of this the

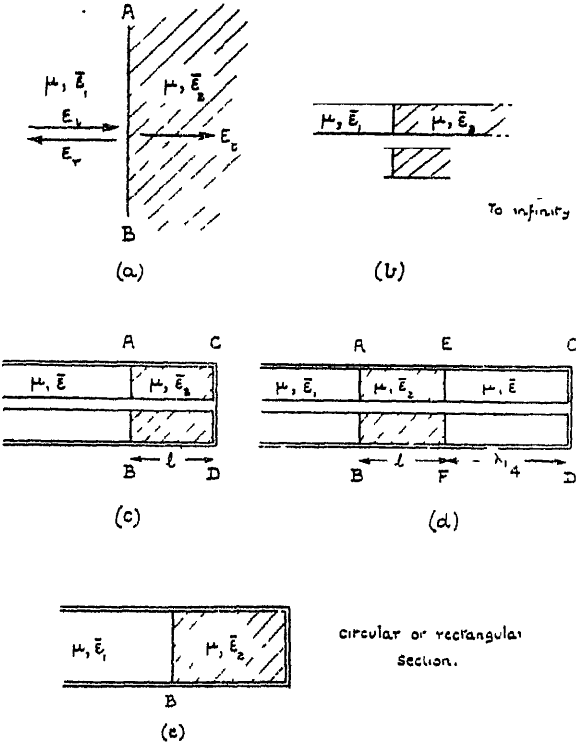


FIG. 5.

magnitude and phase of the reflection coefficient and therefore of $\bar\epsilon_1/\epsilon_2$ can be determined. Thus K is given by the relation

$$K = \frac{\rho - 1}{\rho + 1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

where $\rho = E_{\max}/E_{\min}$ is the Standing Wave Ratio, while

$$\psi = \frac{4\pi}{\lambda_1} \, . \, y' \quad \text{or} \quad \frac{4\pi}{\lambda_1} \, . \, y'' - \pi \quad . \quad . \quad . \quad . \quad (16)$$

where y' and y'' are the distances from the interface of the first maximum and first minimum of electric field intensity respectively, and λ_1 is the wavelength in medium 1.[3]

---

[3] *High Frequency Transmission Lines*, Methuen Monograph, 1945, pp. 76-78.

These relations apply equally if the electromagnetic field is restricted to the enclosure formed by a coaxial transmission line as illustrated in Fig. 5(*b*).

The provision of a semi-infinite length of medium 2, or equivalently of a sufficient length of this medium to ensure complete absorption of the wave transmitted into it, is not of course normally attainable, and the practical procedure in the laboratory is to terminate a limited length of the medium—the test material—with a perfectly reflecting, metallic, sheet C.D. in the manner shown in Fig. 5(*c*). This means that there occurs at the two boundaries AB and CD of the material a series of wave reflections which modifies the effective reflection coefficient of the dielectric interface as viewed from medium 1, and also the standing wave pattern set up in this medium.

Considerable analytical simplification now results from defining for each medium, under conditions of undisturbed wave propagation in them, a characteristic (wave) impedance z_0 equal to the ratio of the transverse electric and magnetic field components.[*] For the two media this characteristic impedance has values $z_{01} = \sqrt{\mu/\epsilon_1}$ and $z_{02} = \sqrt{\mu/\epsilon_2}$ respectively.

Under the semi-infinite conditions illustrated in Figs 5(*a*) and (*b*), the impedance of medium 2 at the interface AB is z_{02} and the reflection coefficient of this boundary for the electric field is given by

$$k = \frac{z_{02}/z_{01} - 1}{z_{02}/z_{01} + 1}$$

which is necessarily identical to (14).

For the condition of Fig. 5(*c*), however, the impedance at the interface has the value $z_{02} \,.\, \tanh \gamma_2 l$,[4] where γ_2 is the propagation constant of the second medium. This leads to a reflection coefficient

$$k = \frac{z_{02}/z_{01} \,.\, \tanh \gamma_2 l - 1}{z_{02}/z_{01} \,.\, \tanh \gamma_2 l + 1}$$

from which it follows that

$$z_{02}/z_{01} \,.\, \tanh \gamma_2 l = \frac{1 + k}{1 - k} \qquad (17)$$

But following Section 1, $\gamma_2 = j\omega\sqrt{\mu\bar{\epsilon}_2}$, and similarly $\gamma_1 = j\omega\sqrt{\mu\bar{\epsilon}_1}$, so that $z_{02}/z_{01} = \sqrt{\bar{\epsilon}_1/\bar{\epsilon}_2} = \gamma_1/\gamma_2$ and (17) may be rewritten

$$\gamma_1 l \,.\, \frac{\tanh \gamma_2 l}{\gamma_2 l} = \frac{1 + k}{1 - k} \qquad (18)$$

k is determinate in magnitude and phase as before from the standing wave pattern in medium 1, and, since this medium will be air, $\bar{\epsilon}_1 (= \epsilon_v)$ and γ_1 are known, and the equation can be solved for γ_2. This, then, affords values for α_2, β_2, ϵ'_2, ϵ_2'', etc., in the manner discussed in Section 1.

An alternative evaluation of the Standing Wave pattern may be used to give $z_{02}/z_{01} \,.\, \tanh \gamma_2 l$ directly. Thus, it may be shown that this impedance ratio is given in magnitude by $E_{\lambda_1/2}/E_{\lambda_1/4}$ where $E_{\lambda_1/2}$ and $E_{\lambda_1/4}$ are the electric field intensities at distances of $\lambda_1/2$ and $\lambda_1/4$ respectively from the interface, and in phase by $\cos^{-1} \dfrac{E_{\max} \,.\, E_{\min}}{E_{\lambda_1/2} \,.\, E_{\lambda_1/4}}$.

A further possible procedure consists in interposing a quarter wavelength of air line between the back face of the test specimen and the terminating metal plate in the manner shown in Fig. 5(*d*). Under these circumstances the impedance presented at the interface AB is $z_{02} \,.\, \cosh \gamma_2 l$, from measurement of which, either directly as mentioned above or via the reflection coefficient, γ_2 is similarly determinate.[5]

---

[4] *High Frequency Transmission Lines*, p. 93.

[5] For still further variations, see *High Frequency Transmission Lines*, pp. 114 and 142.

[*] It will be remembered that in the principal mode of transmission line operation the electromagnetic wave is purely transverse.

The same general principles apply to the use of the analogous wave-guide system illustrated in Fig. 5(*e*), though, since the propagation constant is not so simple a function of the complex dielectric constant, the calculations are rather more involved. Moreover the wavelength range which can be covered by any particular assembly is limited by the cut-off wavelength associated with the size of guide and mode of propagation employed, whereas with the coaxial line operating in the *principal* mode there is no wavelength restriction. The ohmic loss in waveguides is, however, less than in coaxial lines which is an advantage in measurements on low loss materials, though the effect of this loss can be calculated and allowed for.

Conclusion.

The standing wave method of measurement has been used much more widely in the United States than in this country, and Prof. von Hippel and his colleagues at the Radiation Laboratory, Massachusetts Institute of Technology, have developed a highly refined form of coaxial line equipment with which accurate measurements on low-loss materials are possible at wavelengths around 10 cm. They have also produced a series of charts permitting ready evaluation of γ_2 from equation (18). It would seem, however, that in the absence of very great care in the mechanical and electrical design of the standing-wave detector system the accuracy with which the properties of low-loss materials can be measured by this method is not equivalent to that attainable by resonance methods. For tan δ values greater than say 0·01 on the other hand, and also for high permittivity materials, the specimen dimensions suitable for the latter tend to become inconveniently small and the standing-wave method has unquestionable advantages. The wavelength flexibility of the coaxial line arrangement referred to above as compared with the restricted wavelength range of particular resonators is of course a great practical convenience.

Appendix.

Professor Willis Jackson (*Manchester*) said: My discussion of the parameters used for the representation of dielectric behaviour would have been more directly applicable to several of the other papers submitted if I had retained the c.g.s. system of units and had extended my analysis somewhat to cover propagation in wave-guides. I will first make these changes and then seek to correlate the different symbols and nomenclature used by the authors of the papers dealing with measurements at centimetre wave-lengths.

My expression for the propagation constant

$$\gamma = \alpha + j\beta = j\omega\sqrt{\mu\bar{\epsilon}} = j\omega\sqrt{\mu\epsilon_v} \,.\, \sqrt{\frac{\bar{\epsilon}}{\epsilon_v}}$$

may be transformed to its equivalent in the c.g.s. system by noting that

$$\mu\epsilon_v = 4\pi \times 10^{-7} \times \frac{1}{36\pi \times 10^{9}} = \frac{1}{9 \times 10^{16}} = \frac{1}{c^2},$$

where c is the velocity of wave propagation in free space expressed in metres per second, and that the complex relative permittivity or dielectric constant $\frac{\bar{\epsilon}}{\epsilon_v}$ on the m.k.s. system corresponds to the actual permittivity value $\bar{\epsilon}$ on the c.g.s. system, in which ϵ_v has the value unity.* On the latter system therefore

$$\gamma = j\frac{\omega}{c} \,.\, \sqrt{\bar{\epsilon}} = j\frac{2\pi}{\lambda} \,.\, \sqrt{\bar{\epsilon}} = \sqrt{-\left(\frac{2\pi}{\lambda}\right)^2 \,.\, \bar{\epsilon}},$$

---

* In C. G. S. units $\mu = 1$ likewise and therefore need not appear explicitly.

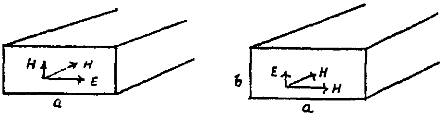
where γ is given in the same linear unit as the free space wave-length λ, that is in cm. units if λ is expressed in cms.

The extension of this expression to cover wave-guide propagation is afforded by equation (1) of the paper by Hall, Halliday, Johnson and Walker, namely

$$\gamma = \alpha + j\beta = \frac{2\pi}{\lambda}\sqrt{\left(\frac{\lambda}{\lambda_c}\right)^2 - \bar{\epsilon}} = \sqrt{-\left(\frac{2\pi}{\lambda}\right)^2\bar{\epsilon} + \left(\frac{2\pi}{\lambda_c}\right)^2} \quad . \quad (a)$$

This differs from the preceding in containing the term $\left(\frac{2\pi}{\lambda_c}\right)^2$, where λ_c is the cut-off wave-length of the particular wave-guide system under consideration when air-filled. The reason for the absence of this term from my expression is that λ_c is infinite for both free-space propagation and for transmission lines propagating in the purely transverse or principal mode. In both these cases there is in principle no restriction on the wave-length at which propagation is possible.

In wave-guides on the other hand λ_c has a finite value depending on the guide dimensions and on the mode type it is desired to propagate. Thus the cut-off wave-length for the H_0 mode in circular section guide, characterised by a circumferential component of electric field and by radial and longitudinal components of magnetic field, is equal to 1·64 times the guide radius. At wave-lengths in excess of this value the H_0 mode cannot be freely propagated. Similarly, for the H_{01} and H_{10} modes in rectangular section wave-guide, each characterised by transverse components of electric and magnetic field and a longitudinal component of the latter, λ_c has values 2b and 2a respectively.



H_{01} mode. FIG. 6. H_{10} mode.

Whiffen and Thompson use somewhat different parameters from those so far described. On page 116 of their first paper they define an attenuation coefficient ψ which differs from my α in that it relates to power and not to the electric or magnetic field intensity. Since, however, the power available at any point in the wave-guide is proportional to the product of the latter intensities, it follows that ψ is equal to 2α.

They also define a propagation constant for the H_{01} mode in rectangular section wave-guide

$$\beta^2 = \epsilon\left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{\pi}{b}\right)^2 = \epsilon\left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{2\pi}{\lambda_c}\right)^2 \quad . \quad . \quad . \quad (b)$$

since for this mode, as illustrated above, $\lambda_c = 2b$. A comparison of this relation with (*a*) above shows that ϵ corresponds to the real part ϵ' of the complex dielectric constant $\bar{\epsilon}$ and that Whiffen and Thompson's propagation constant β corresponds to the phase constant β used by Hall etc. and myself, subject, however, to the normally justifiable condition that α^2 is negligible compared with β^2.

The relation between the attenuation coefficient α and the phase constant β for wave-guide systems and the tan δ value of the dielectric medium which corresponds to my equation (7) for free-space and transmission line propagation is as follows

$$\tan\delta = \frac{2\alpha\beta}{\beta^2 + \left(\frac{2\pi}{\lambda_c}\right)^2 - \alpha^2} . \quad . \quad . \quad . \quad (c)$$

This is the same as the relation given by Whiffen and Thompson subject to the approximation, $\alpha^2 \ll \beta^2$, mentioned above. It may be written, for convenience,

$$\tan\delta = \frac{\alpha\lambda}{\pi\epsilon'} \cdot \sqrt{1 - \frac{1}{\epsilon'} \cdot \left(\frac{\lambda}{\lambda_c}\right)^2} .$$

Collie, Ritson and Hasted express their results in terms of a propagation constant $p_1 = n_1 - j\kappa_1$ which is related to the complex dielectric constant $\bar{\epsilon}$ for the H_{10} mode under consideration by

$$p_1^2 = (n_1 - j\kappa_1)^2 = \bar{\epsilon} - \frac{k_1^2c^2}{\omega^2} = \bar{\epsilon} - \left(\frac{\lambda}{\lambda_c}\right)^2 \quad . \quad . \quad (d)$$

the quantity κ for this mode being $\frac{\pi}{a_1} = \frac{2\pi}{\lambda_c}$. A comparison of this relation with (a) shows that the propagation constant γ as defined by Hall, etc., and myself is related to p by

$$\gamma = j\left(\frac{2\pi}{\lambda}\right) \cdot p_1.$$

Collie's p is my complex index of refraction $\bar{n} = n(1 - j\kappa)$.

Electrotechnics Department,
Manchester University.

DIELECTRIC ABSORPTION IN BENZENE AND LIQUID PARAFFIN SOLUTIONS AT ULTRA HIGH FREQUENCIES.

BY WILLIS JACKSON AND J. G. POWLES.

Received 4th March, 1946.

I. Introduction.

It is well known that solutions of polar molecules in simple non-polar solvents may be expected to exhibit marked dielectric absorption at frequencies of the order of 10^{10} cps. Methods are now available[1] permitting accurate dielectric loss measurements up to and somewhat beyond this frequency, and these methods have been employed for a preliminary study of the variation of absorption with frequency of dilute solutions of benzophenone, nitrobenzene, brombenzene, chloroform and acetone in benzene. For comparative purposes measurements have also been made on solutions of benzophenone in liquid paraffin.

Methods of Measurement.—Table I gives particulars of the resonant systems adopted for measurements at various spot frequencies within the range 10^5 to $2{\cdot}5 \times 10^{10}$ cps.

TABLE I.

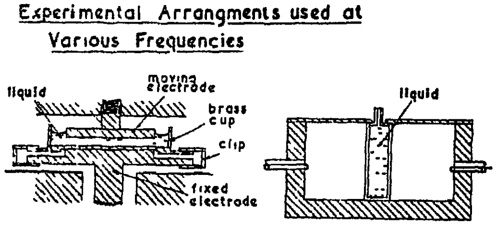
Frequency of Measurement (cps.).	Free-space Wavelength, λ_0 in cm.	Resonant System Employed.	Resonator Dimensions (cm.).	
			Diameter.	Actual Length.
$10^5 - 10^8$	$3 \times 10^5 - 3 \times 10^2$	Lumped * circuit	—	—
$6{\cdot}6 \times 10^8$	45·5	E_{010}	34·66	15·2
$1{\cdot}29 \times 10^9$	23·2	E_{010}	17·78	10·2
$3{\cdot}21 \times 10^9$	9·35	E_{010}	7·15	3·77
$3{\cdot}49 \times 10^9$	8·60	E_{010}	6·58	3·77
✓$9{\cdot}80 \times 10^9$	3·05	H_{01}	4·74	20 (variable)
$2{\cdot}44 \times 10^{10}$	1·23	H_{01n}	1·96	2·7 (variable)

* Hartshorn and Ward dielectric test set.

---

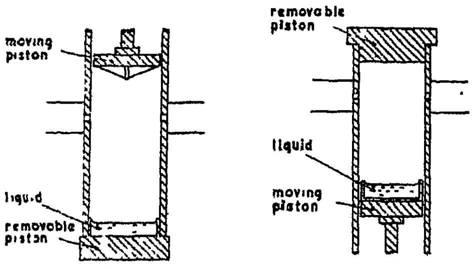
[1] Horner, Taylor, Dunsmuir, Lamb and Jackson, *J.I.E.E.*, 1946, **93**. Part III, 53.

The means adopted for locating the test liquids within the resonant system in the three cases are illustrated in Fig. 1. In the lumped circuit system the liquid was contained in a metal cup clamped to the lower electrode of the variable (air) condenser in the manner shown in (*a*). The presence of the rim of this cup introduced additional edge capacitance, but this could be allowed for without difficulty. The cylindrical form of specimen required by the E_{010} system was provided by enclosing the liquid under test in a thin-walled quartz bottle as illustrated in (*b*). The presence of the bottle wall complicates the dielectric arrangement within the cavity as compared with the case of wholly solid specimens, but mathematical expressions permitting calculation of the relative permittivity and tan δ of the enclosed liquid from measured values of resonant wavelength and *Q*-factor have been derived, and are to be published elsewhere.[2] These calculations involve a knowledge of the dielectric properties of the bottle material and these have been obtained from measurements, by the same technique, on rod specimens of the material.



(a)

(b) E_{01} Resonator



(c) H_{01} Resonator at 9·8 × 10⁹ cps.

(d) H_{01} Resonator at 2·44 × 10¹⁰ cps.

FIG. 1.

With both the lumped circuit and E_{010} systems the values of relative permittivity obtained for a range of non-polar liquids (normal hexane, cyclohexane, benzene, carbon tetrachloride and carbon disulphide) have agreed with the accepted (audio-frequency) values to within ± 1 %.

In the H_{01n} system as used at 9·8 × 10⁹ cps. the liquid was contained in a thin-walled recess machined out of the fixed, but removable, end of the resonator in the manner shown in (*c*). This permitted the liquid to be changed without disturbing the movable tuning piston. The rim of the recess reduced the diameter somewhat at the lower end of the resonator and led to derived permittivity values which were a few per cent. low. The effect cannot conveniently be allowed for analytically, but measurements with several rim thicknesses have enabled an extrapolation to zero wall thickness to be made resulting in permittivity values for the non-polar liquids mentioned above in good agreement with the lower frequency values. The observed tan δ for a given solution was found to be independent of the rim thickness within the accuracy of measurement, and the values obtained with the thinnest practicable rim thickness have been taken as correct. The presence of the meniscus was shown to have no significant effect on the relative permittivity and tan δ values by measurements made with a range of depths of liquid. In the 2·44 × 10¹⁰ cps. H_{01n} resonator [*] the liquid container consisted of a thin-walled metal cup seated on, and moving with, the tuning piston in the manner shown in (*d*). In this case the derived permittivity values have been consistently about 5 per cent. high. The results given in Table II for a sample of polythene show the effect to be due to the presence of the cup. These results suggest that the tan δ values obtained for the test liquids are likely to be some 5 per cent. high, though this lies within the probable accuracy of tan δ measurement.

---

[2] Dunsmuir and Powles, *Phil. Mag.* (*awaiting publication*).

[*] A detailed description of the experimental assembly used at this frequency is to be published by J. Lamb.

TABLE II.

	Derived Relative Permittivity.	Derived tan δ Value × 10^4.
Polythene specimen placed in, and filling the cup	2·39	2·18
Cup removed—polythene specimen alone	2·32	2·14
Cup removed—polythene specimen of same diameter as resonator .	2·31	2·09

The absolute accuracy of the tan δ determinations is not easy to assess since no definite standards of reference are available, but consideration of the possible sources of error involved in the various techniques has led to the estimates given in Table III for the several frequencies of measurement.

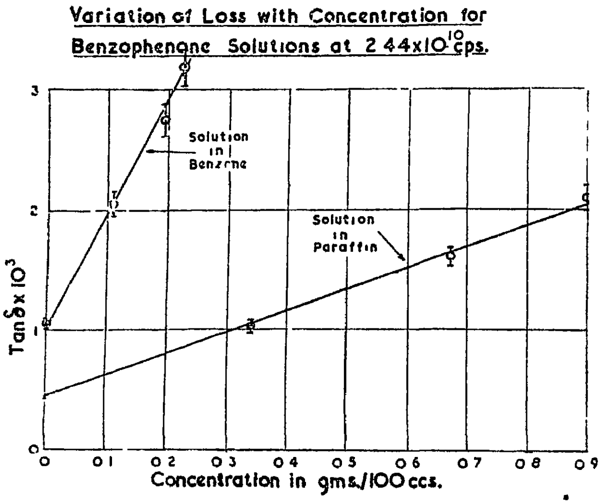
TABLE III.

ESTIMATES OF ACCURACY OF TAN δ DETERMINATION AT THE SEVERAL FREQUENCIES OF MEASUREMENT.

Frequency cps.	$2{\cdot}44 \times 10^{10}$	$9{\cdot}8 \times 10^9$	$3{\cdot}3 \times 10^9$	$1{\cdot}19 \times 10^9$	$6{\cdot}18 \times 10^8$	$8{\cdot}6 \times 10^7$ to $1{\cdot}3 \times 10^5$
Estimated accuracy	±5%			±10%		±5%

II. Statement and Discussion of the Dielectric Loss Results.

(1) Solutions in Benzene.—As stated in the introduction, measurements have been made on solutions of benzophenone, nitrobenzene, brombenzene, chloroform and acetone in benzene. Extreme purity of these polar substances was not considered necessary in view of the limited accuracy of measurement, and the substances used were of "analytical reagent" standard. Analar benzene was employed as solvent throughout the measurements without additional purification, notwithstanding that it exhibited a substantial loss at the highest frequency of measurement. Thus its tan δ value rose from $1{\cdot}4 \times 10^{-4}$ at $1{\cdot}29 \times 10^9$ cps. to $11{\cdot}0 \times 10^{-4}$ at $2{\cdot}44 \times 10^{10}$ cps. The presence of this basic loss has not interfered, however, with the determination of the additional loss—which it is convenient to refer to as the incremental tan δ—associated with the introduction into the benzene solvent of a known concentration of the polar substance under investigation. What is significant to this determination is the change in *Q*-factor of the resonant system which results on replacement of the solvent alone by the solution under test, a change which can be made adequate for accurate incremental tan δ measurement by appropriate choice of the concentration of polar solute consistent with the condition that the concentration is low enough to avoid association effects. Measurements with variable concentration showed the tan δ value to be a linear function of the concentration for any particular solution at each of the frequencies at which measurements have been made, Typical results for benzophenone-benzene solutions are shown in Fig. 2.



n.b. ⏀ indicates limits of assessed accuracy of measurement

FIG. 2.

This linearity permitted a simple reduction of the various results to a standard concentration—chosen as 1 g. per 100 c.c. of solution—so that the behaviour at the several frequencies could be readily compared.

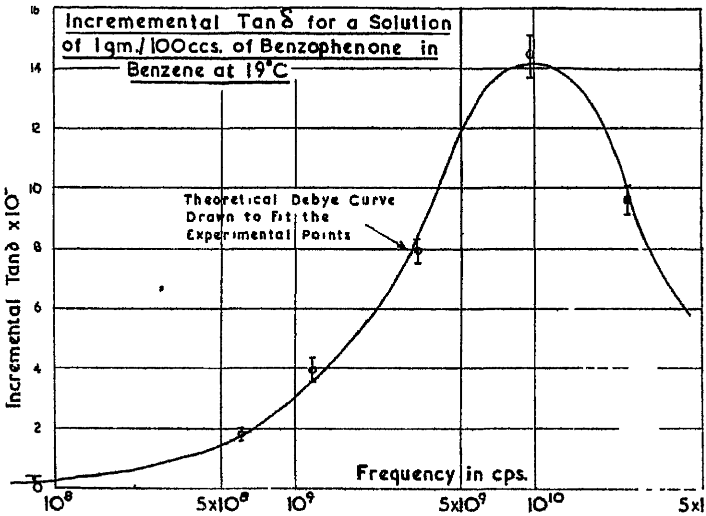


FIG. 3.

The incremental tan δ values obtained at a temperature of 19 ± 1° C. for the five solutions mentioned above are given in Table IV. The benzophenone solution results are plotted in Fig. 3.

Now it is reasonable to expect that the behaviour of these solutions will conform closely to that predicted by the Debye theory since the physical concepts on which this theory is founded would seem to be satisfied as far as is practicable for real polar molecules in liquid solution. This theory gives the variation of tan δ with frequency f cps. of dilute solutions of a single type of polar molecule in a simple non-polar medium in the form

$$\tan \delta = \frac{(\epsilon' + 2)^2}{\epsilon'} \cdot \frac{4\pi}{27} \cdot \frac{\mu^2}{kT} \cdot Nc \cdot \frac{2\pi f \tau}{1 + (2\pi f \tau)^2} \quad . \quad . \quad (1)$$

In this relation ϵ' is the real part of the complex relative permittivity of the solution and, for the solutions used in the present experiments, does not differ significantly from the relative permittivity of the solvent; μ is the dipole moment, and c the concentration in gram molecules per c.c., of the polar solute, and τ is the relaxation time of dipole orientation.

It follows that tan δ passes through a maximum given by

$$\tan \delta_m = \frac{1}{2} \cdot \frac{(\epsilon' + 2)^2}{\epsilon'} \cdot \frac{4\pi}{27} \cdot \frac{\mu^2}{kT} \cdot Nc \quad . \quad . \quad . \quad (2)$$

at the frequency $f_m = 1/2\pi\tau$.

By use of equation (1), the curve of Fig. 3 has been drawn as the closest fit to the experimental points. The agreement is sufficiently close to justify the conclusion that only one relaxation time is involved, and acceptance of this curve as a correct representation of the behaviour of the solution leads to a value of relaxation time for benzophenone in benzene at 19° C. of $16{\cdot}4 \times 10^{-12}$ sec. This agrees closely with the figure of $16{\cdot}1 \times 10^{-12}$ sec. derived by Holzmüller,[3] assuming equation (1), from measurements by a thermal method at 8×10^7 cps. and 24° C.

TABLE IV.

INCREMENTAL TAN δ VALUES ($\times 10^3$) FOR SOLUTIONS IN BENZENE, REDUCED TO A CONCENTRATION OF 1 G. PER 100 C.C.

Polar Solute.	Frequency (cps.).					
	8×10^7.	$6{\cdot}18 \times 10^8$.	$1{\cdot}19 \times 10^9$.	$3{\cdot}3 \times 10^9$.	$9{\cdot}8 \times 10^9$.	$2{\cdot}44 \times 10^{10}$.
Benzophenone .	$< 0{\cdot}4$	1·82	3·95	7·92	14·5	9·65
Nitrobenzene .	—	2·84	7·17	14·1	35·7	29·8
Brombenzene .	—	0·45	0·71	1·52	4·61	4·22
Chloroform . .	—	—	—	0·71	2·42	2·98
Acetone . .	—	—	—	4·4	14·7	29·2

Substitution in equation (2) of the maximum tan δ of the curve of Fig. 3 affords a value of 3·04 Debye units for the dipole moment of the benzophenone molecule, which likewise agrees well with the figure of 2·99 obtained by Muller[4] from refractive index measurements on benzene solutions at audio frequencies.

Similar derivations may be made from the results given in Table IV for the nitrobenzene, brombenzene, chloroform and acetone solutions, though with somewhat less reliability in view of the greater need in these cases for measurements at still higher frequencies. The derived values of dipole moment and relaxation time are given in Table V along with such information as is available from other sources. In the case of the acetone solution, the highest frequency of measurement was not sufficiently close to that of the peak of tan δ to permit determination of the dipole moment, but by accepting the value which has been obtained from audio-frequency measurements, an estimation of the relaxation time can be made.

---

[3] Holzmuller, *Physik. Z.*, 1937, 38, 574.

[4] Muller, *Physik. Z.*, 1933, 34, 689.

TABLE V.

DERIVED VALUES OF DIPOLE MOMENT AND RELAXATION TIME.

Solute Molecule.	Freq. of Maximum tan δ ($\times 10^{10}$ cps.).	Corresponding Wavelength cm.	Dipole Moment μ Debye Units.		Relaxation Time τ $\times 10^{12}$ sec.	
			From Data of Table IV.	From the Literature.	From Data of Table IV, 19 ± 1° C.	From Holzmuller 24° C.
Benzophenone	0·97	3·10	3·04	2·99	16·4	16·1
Nitrobenzene .	1·38	2·18	3·96	3·98	11·6	—
Brombenzene .	1·61	1·98	1·64	1·55	10·5	—
Chloroform .	2·06	1·45	1·13	1·20	7·7	—
Acetone .	4·83	$0{\cdot}62_5$	2·74*	2·74	3·3	2·47

* Assumed value.

2. Solutions in Liquid Paraffin.—Since it appears from the results discussed above that the absorption in benzene solution is characterised in each case by a single relaxation time, it was considered of interest to use a more complex solvent, and measurements were therefore made on solutions of benzophenone in medicinal liquid paraffin.* The benzophenone dissolved readily on slight heating and it was considered that a true solution was formed. A linear relationship between tan δ and concentration was again observed at all frequencies of measurement, which in this case extended downwards to $1{\cdot}3 \times 10^6$ cps. The results obtained, reduced to a concentration of 1 g. per 100 c.c. of solvent are given in Table VI, and are plotted alongside curve (*a*) in Fig. 4. Curve (*b*) is a reproduction of the corresponding benzene solution curve of Fig. 3.

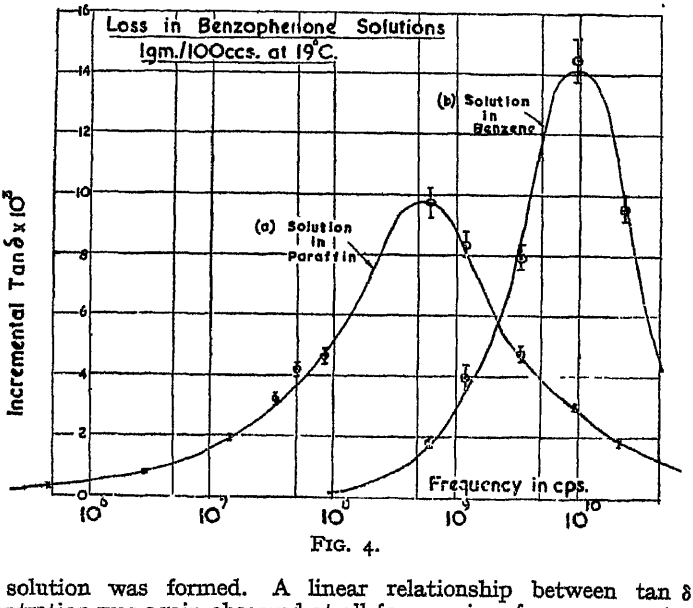


FIG. 4.

---

* Viscosity at 19° C, 1·97 poise.

It is unfortunate that measurements were not possible between 10^8 and 5×10^8 cps. in view of the proximity of the peak tan δ value to this region of frequency. However the symmetrical curve drawn through the available experimental points, and with its peak at $6{\cdot}18 \times 10^8$ cps., probably gives a reliable representation of the overall behaviour.

It will be seen that this curve is broader and its height lower than that for the corresponding benzene solution, indicative of a spectrum of relaxation times. Following Sillars,[5] it is still possible, however, to derive a value of dipole moment, based on the area under the tan δ — log f curve, and since the areas under curves (*a*) and (*b*), Fig. 4, agree to within 3 % * it follows that the dipole moment values for benzophenone afforded by the paraffin and benzene solution results are in agreement within $1\frac{1}{2}$ %.

TABLE VI.

Frequency cps.	$2{\cdot}44 \times 10^{10}$	$9{\cdot}8 \times 10^9$	$3{\cdot}30 \times 10^9$	$1{\cdot}20 \times 10^9$	$6{\cdot}18 \times 10^8$	$8{\cdot}57 \times 10^7$
tan $\delta \times 10^3$	1·98	3·10	4·78	8·35	9·72	4·60
Frequency cps.	$5{\cdot}18 \times 10^7$	$3{\cdot}44 \times 10^7$	$1{\cdot}45 \times 10^7$	$2{\cdot}86 \times 10^6$	$4{\cdot}74 \times 10^5$	$1{\cdot}3 \times 10^5$
tan $\delta \times 10^3$	4·16	3·26	1·88	0·82	0·38	$<0{\cdot}1$

For systems characterised by a spectrum of relaxation times Fuoss and Kirkwood [6] have found it possible to represent the tan δ variation with frequency by the empirical relation

$$\tan \delta_f = A \operatorname{sech} \left(\alpha \,.\, \log_e \frac{f_m}{f}\right)$$

where f_m denotes the frequency at which tan δ attains maximum value and α is a parameter which gives a measure of the breadth of the distribution of relaxation times. For a single relaxation time, as in the case of curve (*b*), Fig. 4, α is unity, and it decreases with increased breadth of the relaxation time spectrum. Curve (*a*), for the paraffin solution, conforms closely to this relation, the appropriate value of α being 0·68.

It is of interest to ascertain whether the relaxation times for the benzophenone molecule in benzene and paraffin are related in an analogous manner to the macroscopic viscosities of these solvents. The desired comparison is made in Table VII based on taking the dominant relaxation time in paraffin as that corresponding to the position of the absorption peak of this solution. It will be seen that the relaxation time and viscosity ratios do not correspond,

TABLE VII.

COMPARISON OF RELAXATION TIME AND MACROSCOPIC VISCOSITY DATA AT 19° C.

(Benzophenone solutions.)

Solvent.	Viscosity η Poise.	Relaxation Time τ $\times 10^{12}$ sec.	$\frac{\eta \text{ par.}}{\eta \text{ ben.}}$.	$\frac{\tau \text{ par.}}{\tau \text{ ben.}}$.
Benzene .	0·0066	16·4	298	18·0
Paraffin .	1·97	295		

---

[5] Sillars, *Proc. Roy. Soc.*, *A*, 1938, **169**, 66.

[6] Fuoss and Kirkwood, *J. Amer. Chem. Soc.*, 1941, **63**, 385.

* There should be a slight difference between these areas on theoretical grounds arising from the factor $(\epsilon' + 2)^2/\epsilon'$ in equation (1), since ϵ' is different for the two solvents, namely, 2·20 for paraffin and 2·28 for benzene. This effect amounts, however, to less than $\frac{1}{2}$ per cent.

from which it may be deduced, as has been done by previous authors, that molecular radii calculated from relaxation time data by use of the relation given by Debye are not of great significance.

Summary and Conclusions.

The development of methods permitting the measurement of dielectric loss at frequencies around 10^{10} cps. has made it possible to determine directly the relaxation time in benzene solution of a number of polar molecules, namely benzophenone, nitrobenzene, brombenzene, chloroform and acetone. It appears that each of these systems is characterised by a single relaxation time, and that the variation of tan δ with frequency conforms to the original Debye theory. This conformity has not previously been demonstrated. Substitution of the maximum tan δ values in the Debye equations affords values of dipole moment for the respective molecules which agree closely with those obtained from audio-frequency refractive index measurements by previous workers.

Measurements on a solution of benzophenone in liquid paraffin have shown the existence in this case of a spectrum of relaxation times. However, by use, following Sillars, of the area under the tan δ—log f curve, these measurements afford a value of dipole moment for the benzophenone molecule which agrees within 1½ % with that obtained from the corresponding benzene solution results. A comparison of the behaviour of the benzophenone-benzene and benzophenone-paraffin solutions confirms the well-recognised inadequacy of the Debye relation connecting the relaxation time of dipole orientation and the macroscopic viscosity of the solvent medium.

The work described forms part of a programme of extra-mural research carried out on behalf of the Ministry of Supply, and the authors express their thanks to the Chief Scientific Officer of the Ministry for permission to publish.

Electrotechnics Department,
The University, Manchester.

SOME MEASUREMENTS OF THE PERMITTIVITY AND POWER FACTOR OF LOW LOSS SOLIDS AT 25,000 Mc./sec. FREQUENCY.*

BY R. P. PENROSE.

Received 20*th March*, 1946.

Introduction.

For measurements on low loss dielectrics near 1 cm. wavelength a resonator method is essential. In a partly filled H_{01} mode cylindrical cavity with transverse discontinuity (Fig. 1) the boundary conditions can be satisfied without introducing waves of other types.[1] Since the cavity can be tuned by varying its length the oscillator frequency can be kept constant.

The H_0 mode of resonance has three important advantages. The electric field near the cylindrical walls is small, so that the specimen need not be a tight fit; no contact is required between the tuning plunger and the walls; finally, the resistive losses are lower than for other modes, so that the H_0 resonator is especially suitable for measurements on low loss materials. The theory of the method is developed in the section which follows, and measurements on 15 solid materials are discussed. These materials were of interest because of their practical application as insulators.

---

[1] Lamont, *Phil. Mag.*, 1940, **29**, 521 ; 1940, **30**, 1.

* Originally described in an Admiralty report (C.L. Misc. 32) dated December 1944.

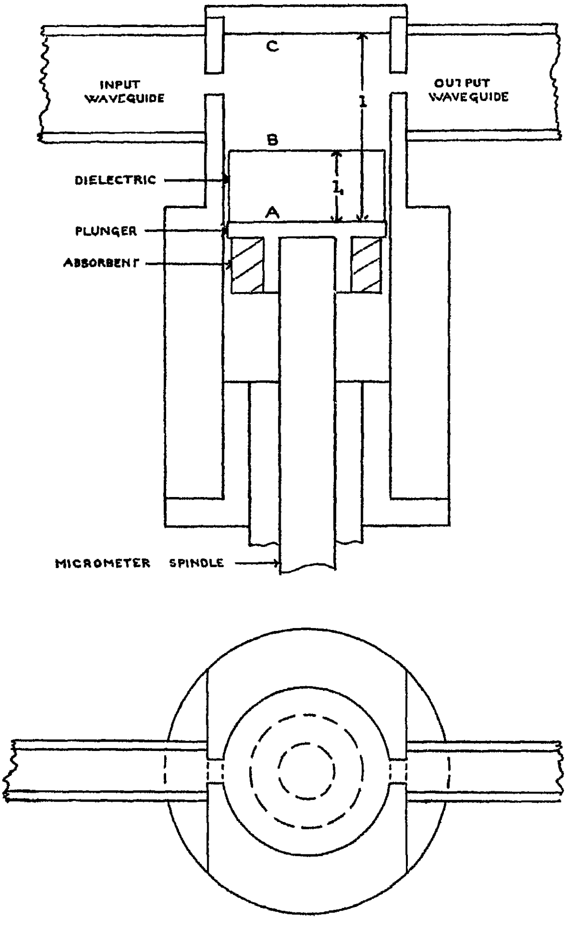


FIG. 1.—Resonator.

Notation.—Field components in the part of the resonator filled with dielectric carry the suffix *d*. Those in the air-space are distinguished by the suffix *t*.

ϵ = permittivity (dielectric constant) of dielectric.
λ = free space wavelength.
$\lambda_c = 2\pi/k$ = critical H_0 wavelength in resonator (radius a).
λ_t = tube wavelength in air space.
λ_d = tube wavelength in dielectric.
z = longitudinal co-ordinate measured from plane of end A of resonator (Fig. 1).
$\omega/2\pi$ = frequency.

$$A = \frac{\text{Electric field amplitude in dielectric}}{\text{Electric field amplitude in air space}}.$$

l_1 = thickness of dielectric specimen.
l_2 = reduction in resonant length of resonator caused by introduction of dielectric.
l = length of resonator.
σ = conductivity of resonator walls.
$\tan\delta$ = power factor of dielectric.
Q = Q-value of resonator containing dielectric.
Q_0 = Q-value of empty resonator.
E_{sd} = total energy stored in dielectric.
E_{st} = total energy stored in air.
E_r = mean power loss on resonator walls.
$n_1 = 4l_1/\lambda_d$.
$n_2 = 4(l - l_1)/\lambda_t$.
ν = number of half-wavelengths in resonator when Q_0 is measured.
dl = change in l required to tune between half-power points when resonator contains dielectric.
dl_0 = corresponding change in l when resonator is empty.

The remaining symbols have their usual meanings. Magnetic quantities are expressed in E.M.U.; electrical ones in E.S.U.

Permittivity.—The introduction of dielectric reduces the length of the cavity at which H_{01} resonance occurs. The boundary conditions at the air-dielectric interface can be satisfied by postulating that at resonance there is an H_{01} wave in the air space and another H_{01} wave in the dielectric. The reduction in cavity length is not simply related to the thickness of the dielectric. Assuming that an H_{01} wave of arbitrary amplitude exists in the air space BC (Fig. 1), we have for its non-vanishing components [2]

$$\begin{aligned} Hz_t &= k^2 J_0(k\rho) \sin 2\pi(z + l_2)/\lambda_t \\ H_{\rho t} &= (2\pi k/\lambda_t)\, J_0'(k\rho) \cos 2\pi(z + l_2)/\lambda_t \\ E_{\phi t} &= (\mathrm{j}\,\omega k/c)\, J_0'(k\rho) \sin 2\pi(z + l_2)/\lambda_t \quad . \quad . \quad (1) \end{aligned}$$

while an H_{01} wave in the dielectric would have

$$\begin{aligned} Hz_d &= Ak^2 J_0(k\rho) \sin 2\pi z/\lambda_d \\ H_{\rho d} &= A(2\pi k/\lambda_d) J_0'(k\rho) \cos 2\pi z/\lambda_d \\ E_{\phi d} &= (\mathrm{j}A\,\omega k/c) J_0'(k\rho) \sin 2\pi z/\lambda_d \quad . \quad . \quad . \quad (2) \end{aligned}$$

A allows for the possible difference in amplitude in the two regions. The continuity conditions at the plane B ($z = l_1$) lead to only two distinct equations, *viz.*,

$$\sin 2\pi(l_2 + l_1)/\lambda_t = A \sin 2\pi l_1/\lambda_d$$
$$(1/\lambda_t) \cos 2\pi(l_2 + l_1)/\lambda_t = (A/\lambda_d) \cos 2\pi l_1/\lambda_d$$

which determine A and l_2. Hence l_2 must satisfy the relation

$$(\lambda_t/2\pi l_1) \tan 2\pi(l_2 + l_1)/\lambda_t = (\lambda_d/2\pi l_1) \tan 2\pi l_1/\lambda_d \quad . \quad . \quad (3)$$

while A is given by

$$A = \left[1 + \left(\frac{\lambda_t^2}{\lambda_d^2} - 1\right) \cos^2 \frac{2\pi l_1}{\lambda_d}\right]^{-\frac{1}{2}} \quad . \quad . \quad . \quad (4)$$

Equation (3) is a particular case of equation (9) in Lamont's first paper.[1] It gives λ_d in terms of the measured quantities λ_t, l_1, l_2 (λ_t is given by the separation of the resonances in the empty cavity). To find λ_d it is necessary to solve an equation of the form $\tan\theta/\theta = \alpha$ for θ, where $\theta \equiv 2\pi l_1/\lambda_d$. A graph of $\tan\theta/\theta$ against θ (or of $\theta/\tan\theta$ against θ when $\theta \approx n\pi/2$) enables one to read off θ directly. The solution is not single-valued, but no difficulty arises if the approximate value of the permittivity is known. Ambiguity could in any case be resolved by repeating the experiment with a different thickness of dielectric.

ϵ is calculated from λ_d by using the relation

$$\epsilon = \lambda^2(1/\lambda_d^2 + 1/\lambda_0^2) \quad . \quad . \quad . \quad . \quad (5)$$

where $1/\lambda^2 = 1/\lambda_t^2 + 1/\lambda_0^2$.

---

[2] Lamont, *Waveguides* (Methuen), p. 76.

The relation between l_1, l_2, λ_t and λ_d is considerably simplified when l_1 is an integral multiple n_1 of $\lambda_d/4$. Then $l_2 = n_1/4\ (\lambda_t - \lambda_d)$. If n_1 is even, l_2 is unaffected by small changes in l_1 since the air-dielectric interface is in a region of weak electric field.

Power Factor.—In a resonator partly filled with dielectric, $1/Q=B \tan \delta$ when resistive losses are neglected. B is that fraction of the total stored energy which is present in the dielectric. B depends not only on the relative volumes of dielectric and air in the cavity, but also on the ratio A of the field amplitudes (equation (4)). The calculation of B can be simplified considerably by restricting n_1 to integral values. Evidently the damping per unit volume of dielectric introduced is large when n_1 is odd and small when n_1 is even. In practice it was not desirable to reduce Q below 500, because insufficient power was then available at the detector. Thus for the higher loss dielectrics (say $\tan \delta = 0{\cdot}01$) it was convenient to make $n_1 = 2$ and $n_2 = 2$ or 4. A dielectric of lower loss ($\tan \delta < 0{\cdot}002$) requires a value of B approaching unity; this can be achieved by increasing n_1 and making it odd. When n_1 is large (*e.g.* 6), however, B is not strongly dependent on whether n_1 is odd or even. The other advantages of a disc an integral number of half wavelengths thick were felt sufficient to warrant its use for all the values of $\tan \delta$ encountered.

The losses on the resonator walls contributed usually between 5 and 10 % of the total loss, and allowance had accordingly to be made for them. The introduction of a dielectric altered considerably the surface current distribution, so that the resistive loss was not the same as when the Q value (Q_0) of the empty resonator was measured. As the walls were silver plated under the same conditions, it seemed justifiable to assume that the surface conductivity σ had a uniform value which could be calculated in terms of Q_0. By definition,

$$Q = \frac{\omega(E_{sd} + E_{st})}{E_r + \omega E_{sd} \tan \delta}.$$

For a disc an even number of quarter wavelengths thick, evaluation of the stored energy and power loss in terms of the field components gives

$$\tan \delta = \frac{1}{Q}\left[1 + \frac{n_2\lambda_t^3}{n_1\epsilon\lambda_d^3}\right] - \frac{4\pi^2C^2}{n_1\omega^3\epsilon}\left(\frac{\omega}{2\pi\sigma}\right)^{\frac{1}{2}}\left[\frac{8}{\lambda_d^3} + \frac{n_1 + n_2\dfrac{\lambda_t^3}{\lambda_d^3}}{a\lambda_0^2}\right]$$

while
$$\frac{1}{Q_0} = \frac{4\pi^2C^3}{\omega^3}\left(\frac{\omega}{2\pi\sigma}\right)^{\frac{1}{2}}\left[\frac{4}{\nu\lambda_t^3} + \frac{1}{a\lambda_0^2}\right].$$

Eliminating σ between these equations

$$\tan \delta = \frac{1}{Q}\left[1 + \frac{n_2\lambda_t^3}{n_1\epsilon\lambda_d^3}\right] - \frac{1}{Q_0 n_1\epsilon}\,\frac{\left[\dfrac{8a\lambda_0^2}{\lambda_d^3} + n_1 + n_2\dfrac{\lambda_t^3}{\lambda_d^3}\right]}{\left[\dfrac{4a\lambda_0^2}{\nu\lambda_t^3} + 1\right]} \quad . \tag{6}$$

The justification for neglecting coupled loss in this calculation is given in the experimental section. Q was determined experimentally by measuring the change in l needed to take the response from half-power to half-power. To find the change in frequency which would have the same effect, it is necessary to know $\mathrm{d}l/\mathrm{d}\lambda$. Now

$$l = (n_1 + n_2)\frac{\lambda_t}{4} - l_2,$$

$$\therefore\ l_1 + l_2 = l_1 + (n_1 + n_2)\frac{\lambda_t}{4} - l.$$

Substituting for $l_1 + l_2$ in equation (3) and rearranging,

$$(n_1 + n_2)\frac{\pi}{2} + \frac{2\pi}{\lambda_t}(l_1 - l) = \tan^{-1}\left\{\frac{\lambda_d}{\lambda_t}\tan\frac{2\pi l_1}{\lambda_d}\right\}.$$

Differentiating w.r.t. λ and substituting $\frac{2\pi l_1}{\lambda_d} = \frac{n_1\pi}{2}$, $l - l_1 = \frac{n_2\lambda_t}{4}$ we find

$$\frac{dl}{d\lambda} = \frac{n_2}{4}\frac{d\lambda_t}{d\lambda} + \frac{n_1}{4}\frac{d\lambda_d}{d\lambda}.$$

From (5) $\frac{d\lambda_d}{d\lambda} = \frac{\epsilon\lambda_d^3}{\lambda^3}$. Similarly, $\frac{d\lambda_t}{d\lambda} = \frac{\lambda_t^3}{\lambda^3}$. Substituting in (7) and solving for $\frac{dl}{d\lambda}$,

$$\frac{dl}{d\lambda} = \frac{n_2}{4}\frac{\lambda_t^3}{\lambda^3} + \frac{n_1}{4}\frac{\epsilon\lambda_d^3}{\lambda^3}.$$

Hence
$$\frac{1}{Q} = \frac{d\lambda}{\lambda} = \frac{4\lambda^2 dl}{n_2\lambda_t^3 + n_1\epsilon\lambda_d^3}.$$

While
$$\frac{1}{Q_0} = \frac{2dl_0}{\nu\lambda_t}\cdot\frac{\lambda^2}{\lambda_t^2}.$$

Substituting these values of $1/Q$ and $1/Q_0$ in (6)

$$\tan\delta = \frac{4\lambda^2 dl}{n_1\epsilon\lambda_d^3} - \frac{2\lambda^2 dl_0}{n_1\epsilon}\frac{\left[\frac{8a\lambda_c^2}{\lambda_d^3} + n_1 + n_2\frac{\lambda_t^3}{\lambda_d^3}\right]}{[4a\lambda_c^2 + \nu\lambda_t^3]} \quad . \quad . \quad (8)$$

A similar calculation for odd values of n_1 gives

$$\tan\delta = \frac{4\lambda^2 dl}{n_1\epsilon\lambda_d\lambda_t^2} - \frac{2\lambda^2 dl_0\left[\frac{4a\lambda_c^2}{\lambda_d^3} + \frac{4a\lambda_c^2}{\lambda_d\lambda_t^2} + n_1 + n_2\frac{\lambda_t}{\lambda_d}\right]}{n_1\epsilon[4a\lambda_c^2 + \nu\lambda_t^3]} \quad . \quad (9)$$

Experimental.

Power from a Klystron reflexion oscillator * was guided to the resonator along an H_1 rectangular waveguide in which an attenuating pad (5 decibels) was inserted; this minimised "pulling" of the oscillator frequency as the resonator was tuned. A similar waveguide conveyed power from the resonator to the crystal detector. A variable attenuator in this waveguide served to adjust the power incident on the crystal; the rectified current (about 1 μA.) was measured on a D'Arsonval galvanometer. The power fed through the resonator varies as Q^2; introduction of the dielectric may reduce Q from 15,000 to 500, corresponding to a 30 decibel drop in transmitted power. The detector sensitivity was improved by matching the crystal unit into the waveguide with two stubs. These were adjusted for maximum detector signal with at least 10 decibels attenuation present in the waveguide. Over 3 decibels of attenuation were retained during the Q measurement in order to avoid increasing the coupled loss by mismatching the detector.

The resonator was 19·00 mm. in diameter, so that $\lambda_c = \frac{19{\cdot}00}{1{\cdot}2197} = 15{\cdot}58$ mm., Power entered and left it through small coupling holes a quarter wavelength from the fixed end, where the longitudinal magnetic field has a maximum value. The upper end C of the resonator could be removed to admit the dielectric specimen; this was placed at the plunger end of the cavity to avoid altering the electrical position of the coupling holes. . The dielectric was held in contact with the plunger by its own weight. The orientation of the feeding waveguides precluded the excitation of E-waves. At the frequency employed in these experiments, H_1, H_2 and H_3 resonances could occur, as well as H_0. The undesired resonances were to a great extent damped out by a piece of high loss material

---

* Designed and constructed in this laboratory by Dr. D. Roaf.

attached to the back of the plunger. The H_0 mode was undamped, for it requires no radial current on the end plates. This arrangement for discriminating against unwanted modes was less effective when dielectric was present in the resonator, but the H_0 resonances could still be identified from their separation. The 0·25 mm. clearance between the edge of the plunger and the wall should not cause serious field distortion, for the surface current is very small near the plunger periphery. The plunger movement was provided by a Moore and Wright micrometer head graduated in 0·002 mm., which was read to 0·001 mm. Sufficient plunger travel was available to give three resonances when the resonator was empty, *viz.*, at $l = \frac{\lambda_t}{2}$; $l = \lambda_t$; $l = \frac{3\lambda_t}{2}$.

Table of Permittivity and Power Factor Values.
(Frequency 25,000 Mc.)

Probable error specified when it exceeds 1 in the last significant figure.

Material.	Origin.	ε.	tan δ.
Polystyrene (Distrene)	BX Plastics Ltd.	2·55 ± 0·02	0·0008
Polystyrene (Transpex 2)	I.C.I. Plastics Ltd.	2·54 ± 0·02	0·0007
Methyl methacrylate (Perspex)	I.C.I. Plastics Ltd.	2·65 ± 0·02	0·012
Ebonite	Britannia Rubber Co. Ltd.	2·73 ± 0·02	0·0038 ± 0·0003
707 (low loss glass)	Corning	4·00 ± 0·05	0·0024 ± 0·0003
C8 (hard glass)	B.T.H.	4·55 ± 0·05	0·006
L1 (lead glass)	G.E.C.	6·8	0·009
Microscope cover glass	Unknown	6·8	0·018
Kodial glass	Plowden and Thompson, Ltd.	4·85 ± 0·05	0·0085 ± 0·0005
Silica	Thermal Syndicate	3·85 ± 0·04	0·0001
Paraffin wax	Anglo-American Oil Co.	2·26 ± 0·02	0·0001
Expanded polyvinyl formal. (Plastazote grade 2)	Expanded Rubber Co. Ltd.	1·16 ± 0·02	0·0031 ± 0·0003
Frequentite	Steatite and Porcelain Ltd.	5·95 ± 0·05	0·0034 ± 0·0002
Frequentite " R "	Steatite and Porcelain Ltd.	6·00 ± 0·05	0·0034 ± 0·0002
Tempradex	Steatite and Porcelain Ltd.	15·1 ± 0·2	0·0002

The oscillator frequency and amplitude were very stable. The drift which occurred while the specimen was being inserted was considerably less than a megacycle and would not introduce appreciable error into the permittivity value. In the power factor determination the half power points were located several times alternately. In practice neither of the points drifted by more than 0·001 mm. The separation dl was of the order of 0·050 mm.

Since a crystal detector was used to locate the half-power points, the power factor value would be inaccurate if the rectified current were not accurately proportional to the square of the input voltage. The rectifying characteristics of several crystals were checked against a thermocouple. When the crystals had a load resistance of 200 ohms and the rectified current did not exceed a few microamps, the deviations from the square law were so small that they would introduce less than 3 % error in the half-power point separation.

In calculating the correction for the losses on the resonator walls, coupled loss was neglected. This loss can be estimated from the fraction f of the power available from the source which is fed through the resonator at resonance. Equivalent circuit considerations show that the ratio of coupled loss to total loss equals $\sqrt{f}$, provided the detector is matched. Direct measurement of f showed that the coupled loss was about 15 % of the total loss in the empty resonator. The ratio of coupled loss to resistive loss is not markedly different when dielectric is present; thus, since the correcting term did not generally exceed 10 % of tan δ, not more than 1·5 % error should be incurred by neglecting coupled loss.

Summary.

The dielectric properties of fifteen low loss solids have been investigated at 25,000 Mc/s. frequency. A flat circular disc of dielectric was inserted at one end of an H_0 mode cylindrical cavity resonator, the reduction in whose resonant length gave the permittivity (Fig. 1). A partly filled cavity was first used for this purpose by Lamont.[1] In an H_0 resonator the electric field near the cylindrical walls is small enough to justify making the specimen a slightly loose fit so as to facilitate its insertion.

Power factors were deduced from the damping of the resonances; these were delineated by detuning the resonator. The correction for the resistive losses on the walls was minimised by using the H_0 mode.

The accuracy of these experiments may be estimated from the scatter between the results obtained with different thicknesses of disc of a given material. The greatest scatter observed was 2 % for ϵ and 8 % or 0·0002 (whichever is the larger) for tan δ. The error in both cases appeared to be governed by the extent to which the specimens could be made flat and parallel. The effect of machining errors was minimised by arranging that the surfaces of the discs were in regions of weak electric field, *i.e.* that their thickness was approximately an integral number of half wavelengths. Under these conditions the experimental values of ϵ agreed within 1 % while the scatter in tan δ did not exceed 5 % (or 0·0001 if tan $\delta <$ 0·002).

The author wishes to make acknowledgement to Lord Cherwell for the use of the facilities of his laboratory; and to Dr. B. Bleaney and Dr. A. H. Cooke for several helpful discussions. He also wishes to thank the Board of Admiralty, under whose auspices the work was carried out, for permission to publish this paper.

The Clarendon Laboratory,
Oxford.

MEASUREMENTS ON THE ABSORPTION OF MICRO-WAVES. I.

BY D. H. WHIFFEN AND H. W. THOMPSON.

Received 26th March, 1946.

Much is now known about the infra-red absorption spectra of molecules in the region of wavelength 1-25μ, where the absorbed energy is used primarily to induce molecular vibration. Little is known, however, of the spectra at much longer wavelengths where the absorbed energy tends to induce molecular rotation. With gases or vapours, analysis of the fine structure of the rotational spectra may lead to a determination of the moments of inertia as well as to other data. With liquids or solutions on the other hand, fresh influences come into play, and the absorption at the very long wavelengths will be expected to consist of a broad region having a maximum at some particular wavelength. In these cases the tendency of an applied electromagnetic field to orient molecular dipoles is resisted by molecular inertia and by the viscous forces of the environment, resulting in a lag between the application of the field and the rotation, with consequent absorption of energy. This phenomenon leads to anomalous dispersion, the fall of dielectric constant with decreasing wavelength, and the absorption of energy is of particular interest in considering the power loss of dielectrics. Debye first formulated a theoretical basis for the phenomenon in terms of the properties of molecular dipoles, and introduced the idea of a " relaxation time " for the process of molecular rotation. This time of relaxation gives a measure of the rate of restoration of random order after removal of the applied field. Later theories of Fröhlich, Onsager, Fuoss, Eyring, Kauzmann and others have attempted to overcome the defects of the original theory of Debye, and to study the relaxation phenomenon as a rate process.

While the absorption of power at longer radio wavelengths has long been used for the determination of dielectric constants, dipole moments, and for detailed examination of the more general theory, few measurements on the absorption of ultra-short radio waves of about 1 cm. wavelength have yet been recorded. The use of this range has certain advantages, one being that the maximum of the absorption band with many mobile liquids lies in this region, the wavelength of this maximum being connected with the time of relaxation by the relation $\omega\tau = 1$. Recent technical developments have now made measurements in this range possible.

The present work began with the object of extending infra-red measurements on hydrocarbons and related molecules to the region of ultra-short radio waves. It was desirable, however, to examine first a wide variety of molecules, not only to develop the experimental methods, but also in order to test the underlying theory. In the latter connection two approaches are possible. Thus, while the original theory of Debye seems satisfactory for many dilute solutions of polar molecules, it is inadequate for polar liquids themselves, presumably due in part to dipolar interactions. In order to formulate a better theory, we can either explore the deviations by gradually increasing the concentration of solutions or making other similar variations, each in turn, thereby revealing the additional factors which must be taken into account, or we can assemble data on pure liquids and solids in which many complicated factors operate simultaneously, and hope to fit them into hypothetical formulations. Whilst the highly accurate determinations of the power loss of solid polymers, plastics, or pure polar liquids may be necessary for industrial use, it has seemed to us preferable first to use dilute solutions of dipoles in non-polar solvents, and to study the variation of the absorption and time of relaxation with different factors such as molecular shape, viscosity of environment, or temperature, if necessary with somewhat less absolute accuracy. For this purpose the discovery of general trends may be preferable to individual measurements of slightly greater absolute accuracy.

Experimental Method.

The present measurements were made at two wavelengths, 1·27 cm. and 3·26 cm., produced by Klystron oscillators fed from a stabilised power pack and cooled by an air blower. The absorption was measured by two methods. In one, the absorption cell, 5-20 cm. in length, was a piece of rectangular wave guide with mica windows, the H_{10} mode being used, and the power transmitted by the cell when empty and filled being compared. Fixed or variable attenuators were placed in front of, or behind, the absorption cell as required. The wavelength was measured by means of an H_{11n} cavity resonator, or by a co-axial cable wave meter, and the power was detected by a crystal connected directly to a moving coil galvanometer. A standing wave detector could also be fitted into the guide circuit as required. In the second method of measurement, the absorption cell was a cylindrical H_{01n} cavity resonator, which could be filled to a known depth with the liquid and tuned to resonance by means of a vertical piston.

A choice between the two methods is influenced by several considerations. Measurements with the wave guide absorption cell are simple, but the volume of the solution required at 3 cm. is inconveniently large. Measurements with a resonating cavity require only small amounts of the solution, and may be essential for substances having a low loss, but they are more exacting experimentally at the shorter wavelengths. After a number of exploratory measurements it was decided to use the wave guide cell at 1·27 cm. and the cavity resonator at 3·26 cm. It was therefore necessary to show that at any given wavelength the two methods lead to the same value of the absorption (tan δ) for a given material, and to examine possible sources of error in each method. The first results suggested values of tan δ determined using the wave guide cell were rather higher than those obtained with the resonating cavity. The most uncertain feature of the former method is the assumptions made about the reflection coefficients at the front and back interfaces of the absorption cell, and in the latter method a small uncertainty arises in the correction which must be made for the curvature of the liquid surface due to surface tension effects. Test measurements proved, however, that neither of these difficulties was as serious as the imperfections of the crystal as a perfect "square law" detector. It was originally assumed that the crystal current recorded by the galvanometer was proportional to the power received, *i.e.* to the square of the electromagnetic field strength. This relationship was tested by calibrating the crystal current against that produced in a sensitive thermocouple, when it was found that different crystals deviate to different extents from the ideal behaviour. Unfortunately a suitable thermocouple for permanent use was not available, and calibration charts were therefore set up and checked at intervals. When the appropriate corrections had been made to the readings of power in the absorption measurements, the two methods of determination gave values of tan δ differing by not more than 5 %, and often agreeing closely. This extent of agreement is satisfactory for the series of measurements described below.

In the first method using the wave guide absorption cell, the power transmitted by the guide tubes was measured first in the absence of the cell and then with the latter in position, and empty. The transmission of the whole empty cell was thus obtained. By means of a standing wave detector in front of the cell, the power at the maximum and minimum of the standing wave could be determined and hence the percentage reflection at the interfaces, and therefore the percentage transmission of these surfaces. The difference between the transmission of the whole cell and that of the reflecting surfaces gave the actual attenuation of the cell. Similar measurements were then made with a filled cell, and hence the attenuation due to the solution could be deduced. It was assumed in this computation that the attenuation of the cell itself was unchanged in the presence of the absorbing material, and that the reflection coefficients of the front and back faces were the same.

The attenuation coefficient, Ψ, can be defined by the relation,

$$T = \text{Transmission} = \frac{\text{power transmitted}}{\text{power received}} = I/I_0 = e^{-\Psi L},$$

or,

$$\Psi = 1/L \log (1/T).$$

It can be shown that for the H_{10} mode Ψ is related to the loss tangent by

$$\tan \delta = \frac{\beta}{(\pi/b)^2 + \beta^2} \cdot \Psi$$

where β is the propagation constant, and a is the longer dimension of the rectangular guide. Also for this mode, β is given by

$$\beta^2 = \epsilon(2\pi/\lambda)^2 - (\pi/a)^2$$

where λ is the wavelength in free space. It was assumed that modes other than H_{01} could be ignored. Any energy transformed into them would appear to have been absorbed. The fact that the measured loss with cyclohexane was zero indicates that their effect cannot be significant.

The theoretical basis of the method with the cavity resonator has been described by Penrose. It involves the determination of Q, π times the ratio of the energy stored in the cavity to the energy lost per half cycle, and tan δ can be related to the value of Q in the absence and presence of the dielectric. For the calculations it is necessary to know the depth of solution or liquid in the cavity, the total length of the cavity at resonance when empty, the change of the resonant length due to the presence of the dielectric, the width of the resonance peak between points of half power with the empty and filled cavity, and the wavelengths of the radiation in air and in the empty guide.

Results.

For the purpose of quantitative comparisons with different molecules it is necessary to refer the power loss, tan δ, of solutions to some arbitrary, but standard concentration, which is chosen as 1 g. mole per 100 c.c., and the loss is then termed the specific loss tangent. We can write, as above, assuming Beer's law of absorption,

$$\text{Transmission} = T = I/I_0 = e^{-c\mathcal{E}L}$$

so that $-\log T/c$ can be calculated if T is measured for a given concentration c. The extrapolation of the attenuation coefficient to the hypothetical concentration of 1 g. mole per 100 c.c. assumes that $-\log T/c$ is constant for all values of c in the range of concentrations used for the measurements. The validity of Beer's law was therefore first tested for a number of solutions of polar molecules in non-absorbing solvents such as benzene. With solutions of toluene in benzene, the plot of $-\log T$ against concentration was linear up to 100 % concentration; and in other cases such as chloroform or acetone dissolved in benzene the plot was satisfactorily linear up to concentrations much higher than those required to obtain a suitably measurable loss. In other cases of solutes which are more capable of association either with themselves or with the solvent, deviations appeared to arise at the higher concentrations.

In considering the subsequent absorption measurements, it was decided in the first instance to assume the original theory of Debye. This leads to the following expression for the loss tangent:

$$\tan\delta = \frac{(\epsilon + 2)^2}{\epsilon} \cdot \frac{4\pi}{27kT} \cdot \mu^2 \cdot c \cdot N \cdot \frac{\omega\tau}{1 + \omega^2\tau^2}$$

in which ϵ is the dielectric constant of the solution, c is the molar concentration, N is Avogadro's number, μ is the molecular dipole moment of the solute, k is Boltzmann's constant, T the absolute temperature, ω the frequency in radians per second, and τ is the time of relaxation of solute molecules.

The absorption will be a maximum when $\dfrac{\omega\tau}{1 + \omega^2\tau^2} = \theta$ has its maximum value of 0·5. In general, for a given solvent and using dilute solutions of a given solute at a fixed concentration at a fixed temperature, tan δ will be proportional to θ, and there will be two values of ω for a given value of tan δ corresponding to wavelengths on opposite sides of the absorption maximum. In order to determine τ it would be best to determine the wavelength corresponding to the maximum of tan δ, but from only two wavelengths this is impossible. We can, however, proceed as follows. In the circumstances just named, at the two wavelengths ω_1 and ω_2,

$$\frac{\tan \delta_1}{\tan \delta_2} = \frac{\theta_1}{\theta_2}.$$

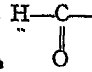
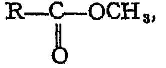
A curve can now be drawn of θ_1/θ_2, called R, against assumed values of τ, from which it is found that for a given value of R there is only one value of τ, but over certain parts of the curve (high and low values of R) the value of τ determined from a measured value of R will not be very accurate. We can therefore determine τ from a measured value of the ratio R with the proviso that the two wavelengths are not far removed from that corresponding to the maximum of absorption. Most of the examples studied could be treated in this way.

When a value has been determined for τ in this way, the dipole moment of the solute can be calculated from the loss at either wavelength. Table I summarises the results of such measurements for a number of solutes dissolved in benzene, and for liquid toluene, which by virtue of its small dipole moment could be regarded as a dilute solution of dipoles in a non-polar medium.

TABLE I.—Temperature 19° C. ± 2° C.

Solute.	Specific Loss tan δ.		τ sec. (Calc.).	μ (Calc.).	μ (Lit.).	τ sec. 25° C. (Lit.).
	1·27 cm.	3·26 cm.				
Benzophenone	1·29	2·19	22×10^{-12}	2·88	2·98	$16{\cdot}1 \times 10^{-12}$
Ethyl benzoate	0·78	0·89	12	1·85	1·82	—
Methyl benzoate	0·67	0·74	12	1·69	1·84	—
Nitrobenzene	3·32	3·58	11·5	3·75	3·9-4·3	—
Bromobenzene	0·58	0·54	10	1·51	1·35-1·58	—
Chlorobenzene	0·63	0·46	7·5	1·52	1·52-1·69	—
Toluene	0·024	0·0175	7·3	0·30	0·35-0·42	—
Cyclohexanone	2·51	1·65	6·5	3·02	2·8	6·78
Nitromethane	1·96	0·91	3·9	2·87	3·02-3·42	—
Acetone	1·43	0·67	3·2	2·58	2·74-2·85	2·47
Methyl acetate	0·57	0·26	3·1	1·65	1·67	—

The first noteworthy result is that the calculated dipole moments agree fairly well with the accepted values, over a wide range. Secondly, the calculated times of relaxation arrange themselves as regards magnitude roughly in accordance with the size of the solute molecule, τ diminishing as the size increases. The times of relaxation given in Table I for 25° C. were measured by Holzmüller[1] at 3·78 metres wavelength, and also fall roughly into line. These results suggest, therefore, that the equations of Debye give a reasonably satisfactory explanation.

Further measurements were then carried out in order to determine how the time of relaxation varies with chain length in such long chain molecules as esters or ketones having a dipole either at the end or in the middle of a hydrocarbon chain, and dissolved in benzene. These solutes differ from those in Table I in being flexible. The loss was measured at each of the two wavelengths for a series of formates  $\mathrm{H{-}\underset{\overset{\|}{O}}{C}{-}OR}$, and for another series of methyl esters of fatty acids  $\mathrm{R{-}\underset{\overset{\|}{O}}{C}{-}OCH_3}$, and values were calculated for μ and τ as described above. The results suggested that for each series, τ increases with increasing chain length of the alkyl radical, and eventually approaches a steady value. This increase would be expected, and the limiting value could be explained by supposing that with long alkyl chains the dipole can rotate more or less freely without much affecting the other end of the chain to which it is attached. Unexpectedly, the results suggested that for an equivalent size of alkyl radical, formates rotate less rapidly than methyl esters, but this might well be a consequence of slight association through the formyl radical.

---

[1] Holzmüller, *Physik. Z.*, 1937, 38, 574.

Although these simple interpretations may be qualitatively correct, however, the quantitative conclusions are less satisfactory. Some inadequacy in the theory is emphasised by the low values calculated for the dipole moments, which depart increasingly from the expected values as the chain length increases. For example, with *n* amyl formate the calculated dipole moment is 1·65, the expected value being 1·90, and with cetyl formate the calculated value is 1·36. Similarly with the methyl esters the calculated values are 1·4–1·5 and the true values about 1·7. These low values found for the dipole moments are the more striking since the dipole enters the expression for the loss as a square, and it therefore seems that the loss cannot vary with ω strictly according to the simple factor $\frac{1}{1+\omega^2\tau^2}$. In this event, the method used for calculating μ and τ will be invalid.

The results therefore suggest that although the equation of Debye can be regarded as roughly satisfactory for rigid solute molecules, some modification is required for the non-rigid solutes.

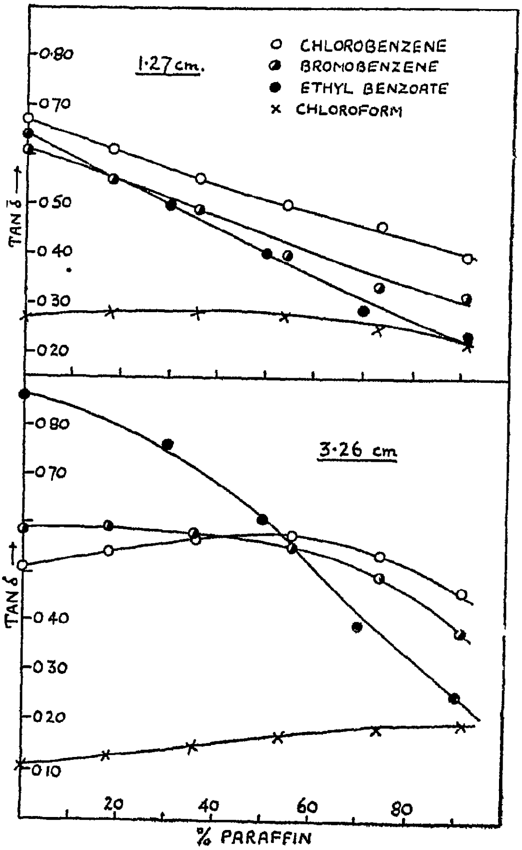


Fig. 1.

Another group of measurements suggests that the simple theory of Debye must not be used without emendation. In this, an examination has been made of the influence of viscosity on the absorption coefficients of a number of rigid solutes in a non-absorbing solvent at each of the two wavelengths and at a fixed temperature. The viscosity was varied by mixing cyclohexane with a high-boiling paraffin, and the measurements and calculations were carried out as already described. Results were obtained at 1·27 cm. and 3·26 cm. using chloroform, bromobenzene, chlorobenzene and ethyl benzoate as solutes. Some other molecules such as acetone and methyl formate were too volatile for convenience, and others like acetonitrile and benzophenone were unfavourable on grounds of insolubility. Measurements at 1·27 cm. were made using the filled wave guide technique, and those at 3·26 cm. using a partially filled H_{01n} resonator. The heavy paraffin used as solvent had a small loss for which a correction was applied.

The results are summarised in Table II, in which the "loss" denotes a specific loss tangent. In computing the specific loss tangent, Beer's law as assumed, as explained above. The significance of A, the "apparent" dipole moment μ, and the "apparent" time of relaxation τ is discussed below.

TABLE II.—TEMPERATURE 19° C. ± 2° C.

% Paraffin.	Loss at 1·27 cm.	Loss at 3·26 cm.	A.	Apparent μ.	Apparent $\tau \times 10^{12}$.
(1) Ethyl Benzoate (Measured in a 10 % solution, in cyclohexane-paraffin).					
0	0·64	0·86	1·72	1·76	16
30	0·50	0·76	1·55	1·68	19
50	0·40	0·61	1·21	1·50	19
70	0·29	0·40	0·80	1·21	18
90	0·24	0·26	0·55	1·00	12
Lit. 1·80 *D*.					
(2) Bromobenzene (Measured in a 4·55 % solution).					
0	0·61	0·59	1·33	1·55	10
18	0·55	0·595	1·28	1·53	12
36	0·49	0·58	1·22	1·49	13
54	0·40	0·56	1·11	1·42	16
73	0·34	0·50	1·06	1·39	19
91	0·32	0·39	0·80	1·20	13
Lit. 1·52 *D*.					
(3) Chlorobenzene (Measured in a 4·55 % solution.)					
0	0·67	0·51	1·37	1·58	8
18	0·61	0·54	1·30	1·54	9
36	0·55	0·57	1·25	1·50	11
54	0·52	0·58	1·24	1·50	12
73	0·46	0·54	1·12	1·43	13
91	0·40	0·47	0·98	1·33	13
Lit. 1·57 *D*.					
(4) Chloroform (Measured in a 4·55 % solution.)					
0	0·27	0·115	0·82	1·22	2·5
18	0·28	0·135	0·65	1·09	4·0
36	0·28	0·15	0·58	1·03	5·0
54	0·28	0·175	0·58	1·02	6·0
73	0·25	0·19	0·51	0·96	7·5
91	0·23	0·20	0·49	0·94	9·0
Lit. 1·18 *D*.					

In Fig. 1 the experimental values for the specific loss tangents are plotted against the composition of the solvent mixture, for each of the two wavelengths. At 3·26 cm. three types of relationship arise, namely: (i) with chloroform, where the loss rises steadily with increasing concentration of paraffin, (ii) with chlorobenzene and bromobenzene where there is a maximum in the curve, and (iii) with ethyl benzoate where the loss falls steadily with increasing concentration of paraffin. According to the expression already given for tan δ, a maximum will be expected at some frequency such that $\frac{\omega\tau}{1 + \omega^2\tau^2}$ is a maximum, namely, 0·5, when $\omega\tau = 1$. At 3·26 cm. the corresponding value of τ is $17·3 \times 10^{-12}$ sec.

If the equations of Debye are applicable, this means that in case (i) τ is less than 17×10^{-12} sec, in case (iii) it is more than this value, and in case (ii) τ will be 17×10^{-12} for the solution corresponding to the maximum. These results would lead us to expect that at the other wavelength, 1·27 cm., the curves for ethyl benzoate, chlorobenzene and bromobenzene will fall in class (iii) and that for chloroform in class (ii). These inferences are confirmed by the measurements.

Qualitatively, the results can therefore be explained reasonably, and the order of relaxation times, namely, chloroform < chlorobenzene < bromobenzene < ethyl benzoate, is that to be expected on the basis of structure.

Once again, however, a quantitative examination of the results reveals serious anomalies. If the equation of Debye is assumed, namely,

$$\tan\delta = \frac{(\epsilon + 2)^2}{\epsilon} \cdot \frac{4\pi}{27kT}\mu^2 c \, . \, N \, . \, \frac{\omega\tau}{1 + \omega^2\tau^2}$$

the only variables in the equation on passing from cyclohexane as solvent to the viscous paraffin are the dielectric constant and time of relaxation. At the most ϵ only changes from 2·0 to 2·1, and the consequential change in the factor $(\epsilon + 2)^2/\epsilon$ can be ignored. We can therefore write,

$$\frac{\tan\delta}{c} = A \cdot \frac{\omega\tau}{1 + \omega^2\tau^2}$$

in which the constant A should depend only upon the dipole moment of the solute, and not upon the composition of the solvent. Values of A and τ required to fit the experimental results at both wavelengths are given in Table II, and the values of μ required to fit the values of A are also included.

It is seen that the values of A are not constant for the same solute, and in some cases τ decreases with increasing percentage of paraffin, that is, with increasing viscosity. Both results are unreasonable, and cannot arise from any plausible experimental error. The trend of A, which falls with increasing viscosity, is the same in each case. The calculated values of the dipole moment for the solutions in pure cyclohexane agree closely with the accepted values, which suggests that the theory is satisfactory for such solutions but breaks down as the long chain paraffin is added.

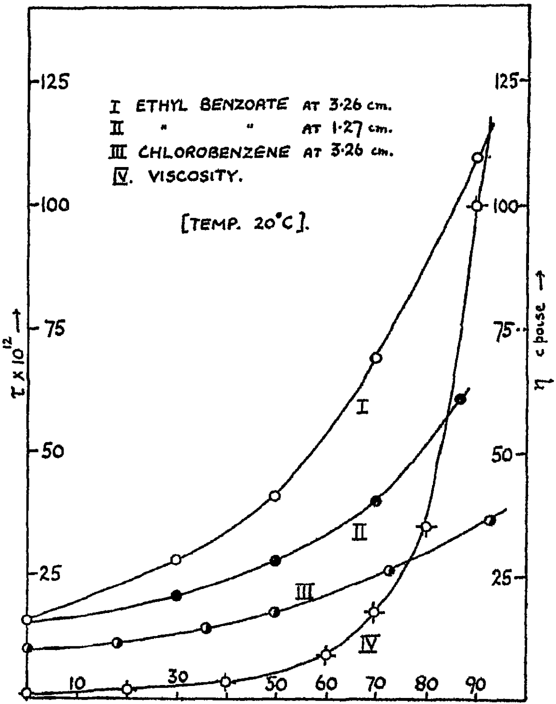


FIG. 2.

Although therefore the times of relaxation calculated in the above way are of doubtful significance, improved estimates can be obtained by assuming A to be constant, with the value found for the solution in cyclohexane, and then calculating τ from the results at each wavelength separately. Although the value of τ thus obtained for a given viscosity depends upon the wavelength used, the difference is not more than about two-fold. Fig. 2 shows the two sets of τ determined in this way for ethyl benzoate, and also the viscosity of the solvent, as a function of the percentage of paraffin. It can be seen that on passing from 0 to 90 % paraffin, τ increases by four to seven-fold, whereas the viscosity increases by some eighty-fold.. Thus the time of relaxation is far from being proportional to the macroscopic viscosity as originally suggested by the equation $\tau = \frac{4\pi\eta a^3}{kT}$, a being the molecular radius. It is again clear, therefore, that η should be regarded as an internal viscosity coefficient.

All these results imply that before any certain inferences can be drawn about the effect of viscosity or inhomogeneity of the medium, or of the flexibility of a solute molecule, upon τ, we must know more precisely how far the equations of Debye are applicable. It is possible that the curves showing tan δ against frequency are broader than suggested by the theory. This can only be discovered satisfactorily from measurements at more than two wavelengths, since the value of ω_{max}—and hence τ—cannot be determined directly from two wavelengths alone. Other wavelengths are now to be used to resolve the difficulty.

In the meantime further information, particularly concerning the variation of relaxation time with temperature, has been obtained. These measurements are described in the succeeding paper.

The above work was sponsored by the Hydrocarbon Research Group of the Institute of Petroleum as an extension of infra-red measurements. We are also greatly indebted to the Telecommunication Research Establishment, Malvern, for apparatus, and to Drs. Bleaney and Penrose for helpful discussions. We also acknowledge gratefully a research assistantship from the Department of Scientific and Industrial Research.

Physical Chemistry Laboratory,
Oxford University.

MEASUREMENTS ON THE ABSORPTION OF MICROWAVES. II.

By D. H. Whiffen and H. W. Thompson.

Received 26th March, 1946.

In continuation of the work described in the preceding paper the variation with temperature of the absorption coefficients of a number of solutions has been measured at wavelengths 1·27 cm. and 3·26 cm. In order to obtain results over a good range of temperature, a non-polar solvent of low freezing-point and high boiling-point is needed in which the solute is sufficiently soluble over the whole range of temperature. Normal heptane was the most suitable. The solutes so far used, namely, chloroform, camphor, methyl benzoate, and α-bromo-naphthalene, were chosen so as to give a wide range of variation of type and shape. Some pure liquids were also examined, chosen to have such a small dipole moment that they could be regarded as dilute solutions of dipoles in a non-polar medium, *e.g.* toluene, *o*-xylene, *p*-cymene and diphenyl methane. These substances have a known dielectric constant, which is almost independent of frequency, and in some cases it was possible to follow the change of absorption above and below the freezing-point.

Experimental.

All the measurements on the effect of temperature were made at wavelength 1·27 cm., using the method of a filled wave guide previously described. The absorption cell was a U-tube immersed in a Dewar vessel, and the required temperatures were obtained with hot water, oil, or baths of solid carbon dioxide-toluene. In order to prevent diffusion of water vapour or toluene into the guide, the couplings near the surface of the bath were sealed with collodion. Mica windows on the guide tube minimised the effect of evaporation at the higher temperature. In order to avoid condensation of water in the cell at the lower temperatures, the extensions to the U-tube were about 15 cm. long, and were not disconnected during a series of measurements, and except during an actual measurement the open ends of the guide were plugged with filter paper.

In the first experiments it was found that non-reproducible results were caused by a change with temperature of the fixed matching attenuator preceding the crystal. An extension was therefore added to the guide so that the attenuator could be attached at a coupling which was at room temperature. The attenuator was made of carbon-loaded plastic which appeared to have a lower temperature coefficient than wood, and it was not attached to hot or cold guide for longer than necessary.

The change from a horizontal straight guide absorption cell to the U-tube meant also that the cell had to be filled by volume. The alternative procedure of entirely filling the cell would necessitate having windows and couplings below the level of the liquid in the thermostat. Moreover, for solutions, it will be sufficient to know the product *cL* of concentration and path length, and this product is independent of temperature. Since the guide is maufactured to highly accurate specifications, its area can be considered as uniform, and the operation of filling by volume should give a satisfactorily accurate measure of the depth. In this argument, the effect of thermal expansion of the copper guide is ignored, although if significant this would affect both the cross-sectional area and the guide propagation constant. For a change of temperature of 100° C., the change in a linear dimension is only 0·15 %, and in cross-section 0·3 %, which can at present be neglected. The liquid-air interfaces in the U-tube cell are distorted by surface tension, and the assumption of equality of reflexion at the two interfaces is therefore not strictly justified, but the results prove that this error cannot be significant.

TABLE I.—6 c.c. in Cell. Wavelength 1·27 cm.

(1).	(2).	(3).	(4).	(5).	(6).	(7).	(8).	(9).	(10).	(11).
−80	0·63	2·6	0·118	0·96	0·0147	0·0141	2·72	39	—	—
−70	0·74	2·5	0·119	0·95	0·0172	0·0166	3·36	30	—	—
−60	0·85	2·5	0·119	0·94	0·0196	0·0191	4·06	23	—	—
−50	0·96	2·5	0·120	0·93	0·0221	0·0218	4·86	19·5	—	—
−40	1·04	2·5	0·120	0·92	0·0236	0·0235	5·47	17	—	—
−30	1·12	2·4	0·121	0·91	0·0253	0·0250	6·07	14·5	—	—
−20	1·18	2·4	0·121	0·90	0·0265	0·0270	6·83	12	—	—
−10	1·20	2·4	0·121	0·89	0·0266	0·0274	7·21	10·5	—	—
0	1·20	2·4	0·122	0·88	0·0265	0·0275	7·51	9·2	0·76	1·21
10	1·18	2·3	0·122	0·87	0·0258	0·0272	7·69	8·5	0·67	1·27
20	1·15	2·3	0·123	0·86	0·0251	0·0268	7·85	7·7	0·59	1·30
30	1·12	2·3	0·124	0·85	0·0243	0·0262	7·94	6·7	0·52	1·29
40	1·07	2·2	0·124	0·84	0·0230	0·0250	7·82	6·0	0·47	1·27
50	1·00	2·2	0·125	0·84	0·0216	0·0236	7·61	5·0	0·42	1·19
60	0·93	2·2	0·125	0·83	0·0200	0·0221	7·36	4·6	0·38	1·21
70	0·88	2·2	0·126	0·82	0·0187	0·0210	7·20	4·3	0·35	1·23
80	0·84	2·1	0·127	0·81	0·0178	0·0202	7·13	4·1	0·32	1·28

The satisfactory working of the method of filling by volume was proved by the close proportionality of the measured loss and the depth of liquid (*i.e.* the volume introduced) for a given solution.

Results.

Table I gives the results for toluene, in the following columns: (1) Temperature, °C.; (2) Log (% transmission), related to tan δ as discussed in the preceding paper; (3) Dielectric constant ϵ, based upon values in the International Critical Tables, adjusted slightly where necessary at lower temperatures to allow for the anomalous dispersion; (4) The ratio tan δ/Ψ, where Ψ is the attenuation coefficient; (5) Density; (6)

tan δ; (7) specific loss tangent, the loss tangent at 1 g. mole per 100 c.c.; (8) $T \times$ specific loss tangent; (9) τ in units 10^{-12} sec.; (10) η, coefficient of macro-viscosity, c. poise ϵ; (11) the ratio $\tau/\eta \times 10^9$.

In Fig. 1 the values of tan δ are plotted against temperature, and show a maximum at about − 10° C.

We can again assume as a basis for discussion the Debye equation,

$$\tan \delta = \frac{(\epsilon + 2)^2}{\epsilon} \cdot \frac{4\pi}{27kT} \cdot \mu^2 \cdot c \cdot N \cdot \frac{\omega\tau}{1 + \omega^2\tau^2}$$

The factor $(\epsilon + 2)^2/\epsilon$ is effectively independent of temperature, changing only from 8·2 to 8·1 over the range of temperature studied. But the expression for tan δ also contains the factor $1/T$, and τ will be strongly dependent upon temperature. The graph of T tan δ/c against temperature is also shown in Fig. 1, having a maximum at about 30° C. At this peak, $\omega\tau = 1$ since $\omega\tau/1 + \omega^2\tau^2$ is a maximum, and the height should be given by $\frac{(\epsilon + 2)^2}{\epsilon} \cdot \frac{4\pi}{27k} \cdot \mu^2 \cdot N \cdot \frac{1}{2}$. The maximum value of T tan δ/c is 7·94, from which the calculated dipole moment μ is 0·32. The value calculated in the previous paper from the loss at 1·27 cm. and 3·26 cm. was 0·30.

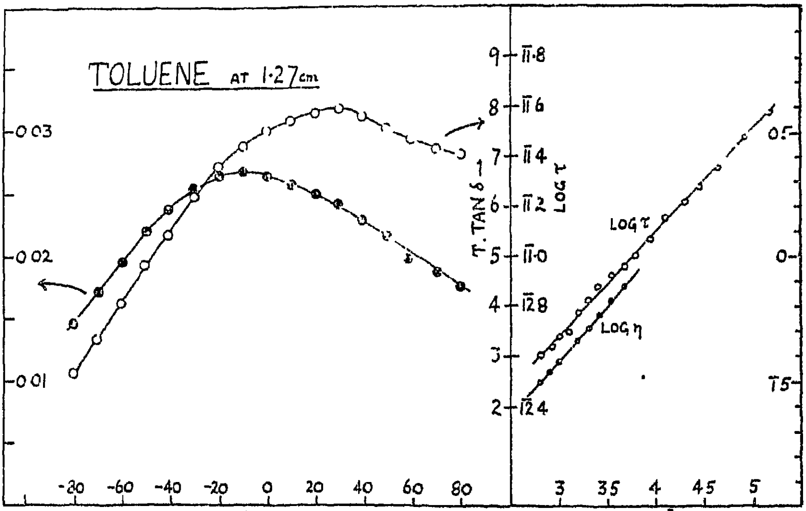


Fig. 1.

The ratio of the value of T tan δ/c at a given temperature to its maximum value will be $2\omega\tau/1 + \omega^2\tau^2$ in which τ is the time of relaxation at the given temperature. The choice between the two possible values of τ for one value of this ratio is determined by whether the temperature is above or below 30° C. (that corresponding to the maximum), at which τ is unambiguously $6{\cdot}7 \times 10^{-12}$ sec. At 19° C. for example, τ is $7{\cdot}7 \times 10^{-12}$, whilst by the independent method of comparing the loss at two wavelengths described in the preceding paper it was found to be $7{\cdot}3 \times 10^{-12}$ sec.

Column (11) in Table I shows that the ratio τ/η is essentially constant. If the original assumption of Debye holds, namely, $\tau = 4\pi\eta a^3/kT$, the volume of each molecule is found to be 16 Å³. If calculated from the density, this volume is 170 Å³. The discrepancy is far greater than could be explained by neglect of free space between packed molecules, and we have to conclude that the simple expression is inadequate. Presumably it is again incorrect to regard η as a macroscopic viscosity.

The relation between viscosity and temperature can be expressed by

$$\eta = A \, . \, e^{E_\eta / RT}$$

in which E_η can be called the molar activation energy of viscosity. Similarly we can write for the time of relaxation,

$$\tau = B \, . \, e^{E_\tau / RT}$$

in which B is first regarded as independent of temperature and E_τ is the activation energy for relaxation. Figure 1 shows plots of log η and log τ against $1/T$. The former gives $E_\eta = 2$ kcal. Since η and τ have been shown above to be roughly proportional, the other plot is expected and found to be roughly linear, with slope about 1·9 kcal. This value for the energy of activation of relaxation may appear small compared with the much higher values (10-50 kcals.) which have been found for some highly polar liquids of solid waxes, glasses and polymers, but in toluene the intermolecular forces will be far weaker.

Similar determinations were made with other liquid hydrocarbons and with solutions of polar molecules in heptane. In the latter case a small correction was made for the very small dielectric loss of the heptane itself. It was also necessary to know the dielectric constant of heptane at different temperatures, and values were estimated from that at 20° C. by assuming that the molecular polarisability $(\epsilon - 1)/(\epsilon + 2) \, . \, M/\rho$ is independent of temperature. A similar uncertainty in ϵ arises with the pure hydrocarbons such as *o*-xylene, particularly in the solid state, but satisfactory values could be estimated.

The results are summarised in Table II.

TABLE II.

Temp. °C.	Specific Loss Tangent.	$\tau_{\text{calc.}} \times 10^{12}$.	
		Chloroform (in heptane).	
−70	0·41	11·5	Solution used was 5 % vol. Dipole moment calculated from maximum of plot is 1·10 ; literature gives 1·18
−60	0·42	9·5	
−50	0·42	8·0	
−40	0·41	6·7	
−30	0·38	5·8	
−20	0·36	5·2	
−10	0·33	4·4	
0	0·30	3·8	
10	0·28	3·6	
20	0·26	3·4	
40	0·23	3·1	
60	0·20	2·7	
80	0·17	2·4	
		Camphor (in heptane).	
−70	1·71	25	Solution used was 2·3 g. per 100 c.c. Dipole moment calculated was 2·98 ; lit. gives 2·95
−60	1·93	21	
−50	2·12	17·5	
[illegible]	2·24	15·5	
[illegible]	2·35	13·5	
[illegible]	2·40	12	
[illegible]	2·48	11	
0	2·50	9·0	
10	2·46	8·0	
20	2·42	7·0	
40	2·24	5·6	
60	2·04	4·8	
80	1·81	4·1	

TABLE II (*cont.*)

Temp. °C.	Specific Loss Tangent.	$\tau_{\text{calc.}} \times 10^{12}$.	
	Methyl Benzoate (in heptane).		
−40	0·63	19	Solution used was 5 % vol. Calculated dipole moment 1·70; lit. gives 1·8-2·0. Solid separates below −40° C.
−30	0·68	16	
−20	0·72	13·5	
−10	0·76	11·5	
0	0·78	9·0	
10	0·79	8·1	
20	0·77	7·5	
40	0·73	6·7	
60	0·68	6·0	
80	0·62	5·1	
	α-Bromo-Naphthalene (in heptane).		
−70	0·23	60	Solution used was 5 % vol. The maximum of the *T* tan δ plot was not quite reached at 80° C. and had to be estimated by extrapolation. This may lead to a small error in the values of τ at the higher temperatures. Dipole moment calculated 1·66; best lit. value 1·58.
−60	0·27	51	
−50	0·31	42	
−40	0·35	34	
−30	0·39	28	
−20	0·42	25	
−10	0·46	22	
0	0·48	20	
10	0·51	18	
20	0·53	16	
40	0·55	13·5	
60	0·57	11·5	
80	0·59	—	
	***p*-Cymene** (liquid).		
−70	0·0049	35	The loss of *p*-cymene is small and would have been better measured in a cavity resonator. Dipole moment calculated 0·19. The only lit. value is 0·29, but is unreliable.
−50	0·0067	22	
−30	0·0080	16	
−10	0·0087	12	
10	0·0090	11	
30	0·0089	8	
50	0·0085	7	
70	0·0081	6·5	
100	0·0073	5·6	
150	0·0061	4·7	
	o-Xylene (liquid and solid).		
−60	0	—	Dipole moment calculated 0·53; lit. value 0·58. The small loss at −30° C. might have been due to incomplete solidification. Melting point −25° C.
−40	0	—	
−30	0·002	—	
−25	0·06	17·5	
−20	0·063	16·0	
−10	0·065	14·5	
0	0·068	12·5	
10	0·069	11·5	
20	0·070	10·3	
40	0·070	8·4	
60	0·068	6·9	
80	0·063	5·9	
100	0·058	5·2	
120	0·053	4·4	
140	0·047	3·8	

Viscosity of Heptane.

Temperature °C.	10	20	40	60	80
η c. poise	0·46	0·41	0·33	0·28	0·23

The plots of tan δ/c against temperature are shown in Fig. 2. The abrupt disappearance of the dielectric loss with *o*-xylene on passing the freezing point is noteworthy, and it must be inferred that rotational freedom in the liquid is completely absent in the solid state. If any rotation in the solid does occur, the time of relaxation must be very high.

TABLE III.

Substance.	From two Wavelengths.		From Temperature Variation.	
	τ.	μ.	τ.	μ.
Toluene	$7{\cdot}3\times10^{-12}$	0·30	$7{\cdot}7\times10^{-12}$	0·32
p-Cymene . . .	$10{\cdot}0\times10^{-12}$	0·19	$9{\cdot}5\times10^{-12}$	0·19
o-Xylene	$8{\cdot}5\times10^{-12}$	0·51	$10{\cdot}5\times10^{-12}$	0·53
Chloroform (in heptane) .	$2{\cdot}4\times10^{-12}$	1·23	$3{\cdot}4\times10^{-12}$	1·10
Camphor (in heptane) .	$6{\cdot}8\times10^{-12}$	2·97	$7{\cdot}1\times10^{-12}$	2·98
α-Bromo-naphthalene (in heptane) . . .	$13{\cdot}5\times10^{-12}$	1·56	$16{\cdot}0\times10^{-12}$	1·66

In this connection we have also measured the loss as a function of temperature for diphenyl methane and tertiary butyl benzene, traversing the melting-points which lie at 25° C. and − 58° C. In these cases the loss, though falling markedly at the freezing-point, did not fall to zero in the solid, and in the solid state there was a roughly linear decrease of tan δ with falling temperature. If this loss in the solid state were real it would imply some form of molecular or intra-molecular rotation. However, although the substances used were adjudged pure, it is still doubtful whether very small amounts of impurities do in fact give rise to the loss.

The values of the dipole moments required to fit the results are in all cases in reasonable agreement with the accepted values. This can be taken as a satisfactory justification for using the equation,

$$\tan\delta = \frac{(\epsilon+2)^2}{\epsilon}\cdot\frac{4\pi}{27kT}\cdot\mu^2\,.\,c\,.\,N\,.\,\frac{\omega\tau}{1+\omega^2\tau^2}$$

in calculating the values of τ. The values of τ thus found also vary for the different molecules in a way which is roughly consistent with their size.

In Table III values are given of μ and τ obtained on the one hand from the measurements at two wavelengths, as described in the preceding paper, and on the other from the temperature variation. The results are in good agreement.

The plots of log τ against $1/T$ shown in Figs. 3*a* and 3*b* are roughly linear, implying the relationship

$$\tau = B\,.\,e^{E_\tau/RT}.$$

TABLE IV.

	kcal.		
	E_τ.	E_η.	
α-Bromo-naphthalene	1·8	2·0	heptane
Methyl benzoate .	1·8	2·0	
Camphor . . .	1·7	2·0	
Chloroform . . .	1·5	2·0	
Toluene . . .	1·9	2·0	
p-cymene . . .	1·7	*ca.* 2·0	
o-xylene . . .	1·9	2·3	

The values of E_τ measured from the slopes of the plots are given in Table IV, which also includes the corresponding values of E_η. In all cases E_τ is approximately equal to E_η, and although the differences are almost within the possible experimental error, it is significant that in no case is E_τ greater

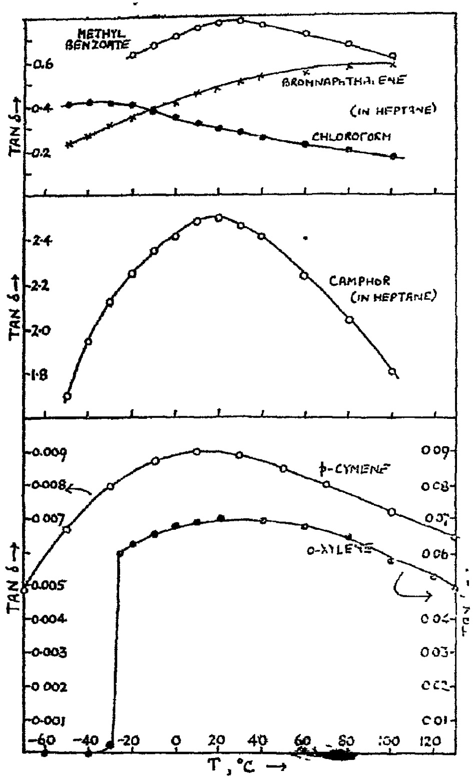


FIG. 2.

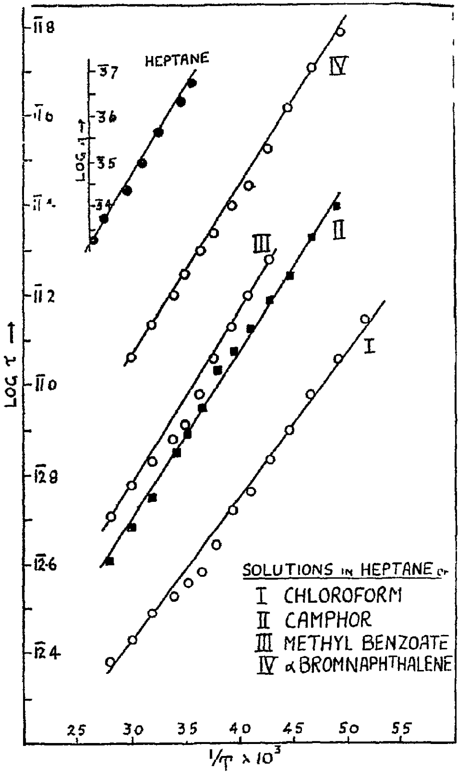


FIG. 3(*a*).

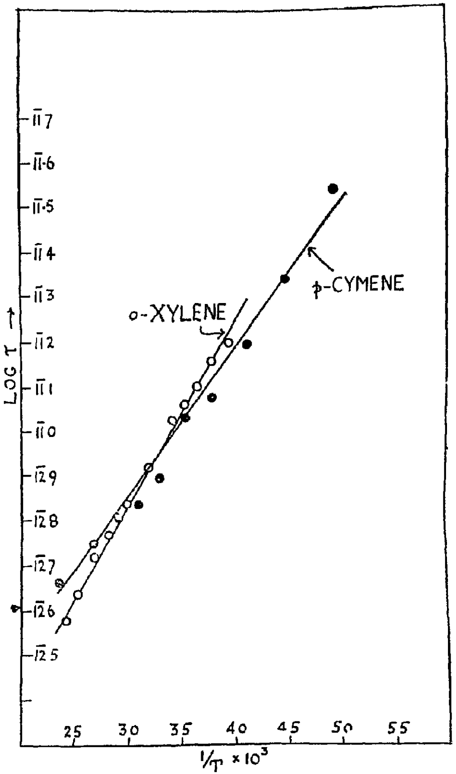


FIG. 3(*b*).

than E_η, and it seems likely that the latter exceeds the former by an amount up to a few hundred calories.

Another approach to the understanding of relaxation as a rate process has been used by Eyring and his co-workers. They write:

$$1/\tau = kT/h \,.\, e^{\Delta S/R} \,.\, e^{\Delta H/RT}$$

where ΔS, ΔH are the entropy change and heat content change for the process of activation. From this equation,

$$\frac{d\,(\log \tau)}{d(1/T)} = \frac{\Delta H + RT}{R} = \frac{E_\tau}{R}.$$

The slope of the plot of log τ against $1/T$ should therefore increase slightly as the temperature is raised. The present results are not sufficiently accurate to detect such curvature. Values of ΔH and ΔS which satisfactorily fit the data are given in Table V. It is noticeable that the entropy changes of activation are all small and negative. The physical significance of such negative values must be left for later consideration, though they are not unknown in other cases,[1] and may be informative about the packing of molecules when compared with the corresponding values for other systems. It seems unlikely, however, that the negative values can arise from experimental errors, and it seems that they are related with the unusually low values of ΔH.

TABLE V.

	E_τ kcal.	ΔH kcal.	ΔS e.u.
α-Bromonaphthalene	1·8	1·2	− 5·1
Methyl benzoate	1·8	1·2	− 3·6
Camphor	1·7	1·1	− 3·8
Chloroform	1·5	0·9	− 3·0
Toluene	1·9	1·3	− 3·2
p-Cymene	1·7	1·1	− 4·2
o-Xylene	1·9	1·3	− 4·0

The above work was sponsored by the Hydrocarbon Research Group of the Institute of Petroleum as an extension of infra-red measurements. We are also greatly indebted to the Telecommunication Research Establishment, Malvern, for apparatus, and to Drs. Bleaney and Penrose for helpful discussions. We also acknowledge gratefully a research assistantship from the Department of Scientific and Industrial Research.

Physical Chemistry Laboratory,
Oxford University.

---

[1] Powell and Eyring, *Advances in Colloid Science*, vol. I, p. 213.

DIELECTRIC PROPERTIES OF WATER.

BY C. H. COLLIE, D. M. RITSON AND J. B. HASTED.

Received 29*th March*, 1946.

The measurement of the dielectric properties of polar liquids at the high frequencies which the development of Radar has made accessible may prove to be a valuable method of studying the liquid state. For this purpose the elementary theory of polar dielectrics needs to be expressed in a form as free as possible from special assumptions about the nature of the dielectric.

The polarisation P of a dielectric will be due to two separate causes so that

$$P = P_1 + P_2 = P_1 + \alpha_1 E'$$

in which P_1 is the dipole moment per unit volume due to the orientation of the molecular dipoles and P_2 is the contribution from the dipole moment induced in the molecules by the field E' acting upon them. The constant α_1 represents the electronic plus atomic polarisability of the molecules. The internal field E' acting upon the molecule will be related to the electric field E of elementary dielectric theory by a relation of the type

$$E' = E + \beta P_1 + \gamma P_2$$

in which β and γ are constants depending on the structure of the liquid for which no completely satisfactory theoretical expressions are available. Since P_2 represents an electronic plus atomic polarisation, γ should not be very different from the $4\pi/3$ originally suggested by Lorentz for β.

Imagine a polarised liquid immediately after the internal polarising field has been destroyed ; molecular collisions will break up the orientation of the molecules and the liquid will return to the unpolarised state at a rate to be measured in terms of a relaxation time $1/k$. Simple arguments show that the shape of the polarisation/time curve cannot be arbitrary if the dielectric is to behave linearly and have properties expressible in terms of a complex dielectric constant $\epsilon_1 - j\epsilon_2$; the most general form of the decay curve is given by the relationship

$$P_1 = A\,e^{-k_1 t} + B\,e^{-k_2 t} + \dots$$

in which there are an indefinite number of relaxation times

$$1/k_1,\ 1/k_2,\ 1/k_3,\ \text{etc.}$$

We shall assume that there is only one relaxation time, $1/k_1$. Whether this assumption is true is a matter which can be determined by experiment, but on theoretical grounds it seems a not unreasonable assumption for a liquid.

With this assumption the polarisation P_1 of a dielectric will obey an equation of the type

$$dP_1/dt = -kP_1 + k'E'$$

in which kP_1 represents the effect of the molecular collisions breaking up the molecular orientation and $k'E'$ the rate at which the dipoles are being orientated by the field E' acting upon them.

For an alternating electric field E will be of the form

$$E = E_0 e^{j\omega t}$$

and P_1 is then given by

$$P_1 = P_{01} e^{j(\omega t + \psi)} \text{ in which } \psi \text{ is a phase constant.}$$

Hence one arrives at an equation of the form

$$\epsilon - \epsilon_0 = \frac{\epsilon_s - \epsilon_0}{1 + j\omega\tau}$$

where ϵ is the dielectric constant, partly real and partly imaginary, at a frequency ω; ϵ_0 is the dielectric constant due to the atomic and electronic polarisation ; ϵ_s is the static dielectric constant; τ is a relaxation time, the relation of which to k and k' is determined by the values assumed for β and γ. The values chosen for these constants are determined by the particular dielectric theory employed. In the case of the Onsager theory, β is not constant, but determined by the value of ϵ.

We are planning a programme of measurements over a wide range of frequency and temperature to investigate the validity of the assumption of a single relaxation time and if possible any dependence of the value of the parameter β on ϵ. In particular it is hoped to obtain useful information by comparing the dielectric properties of water and heavy water.

Experimental.

The methods to be described have been adopted after considerable experimental work on other methods and are capable of at least 1 % accuracy. A satisfactory method of measurement must satisfy the following criteria :

(*a*) it should need only small amounts of liquid in a form capable of exact temperature control.

(*b*) no reliance should be placed on the crystal rectifier obeying an exact law ; at best, crystals are unstable and hard to calibrate.

By working at low input powers and using the conventional superheterodyne technique the difficulties of power measurements are greatly mitigated. The crystal mixer has a linear response over a wide range of input signal and this linear response is unaffected by small changes in the crystal characteristics. The calibration of the instrument is thus transferred to the I.F. for which standard technique is available. Further, owing to the great sensitivity (10^{-12} watts) matching difficulties can be considerably reduced by interposing attenuating pads between each section of the apparatus. This is particularly important at the highest frequencies, as the power output of the valves is low and very sensitive to the impedance offered by the load. In all the work to be described the intermediate frequency was 60 megacycles and all power supplies were stabilised electronically against mains fluctuations.

Measurements at 10 cm.—The first method tried was a modification of the E_{010} resonator method used by Willis Jackson for measuring dielectrics of low power factor. It has the advantage that very small quantities of liquid are needed and that it might be possible to work up to the critical temperature. The resonator used is shown in Fig. 1. The liquid is contained in a glass capillary lying along the axis of the resonator. Three measurements are made : the resonant frequency of the cavity, the resonant frequency with the empty capillary in position and finally the resonant frequency when the capillary is filled with water. The final frequency adjustment is made with the adjustable tuning sleeve. Using a capillary 0·6 mm. in diameter, a change of resonant frequency of 60 mc. was obtained, which could be measured to 1 %. Following the treatment given by Willis Jackson, but allowing for the complex dielectric constant the value of ϵ_1 can be computed from the frequency shift and an approximate value of Q. It is usually sufficient to subtract the shift due to the empty capillary but the full set of equations using three boundary conditions were solved to verify this. If the Q curve of the resonator could be measured equally well, complete data for measuring ϵ_1 and ϵ_2 would be yielded by this method. A really satisfactory method of measuring the resonator damping has not yet been devised ; there appears to be no reason why the difficulties encountered should not be overcome.

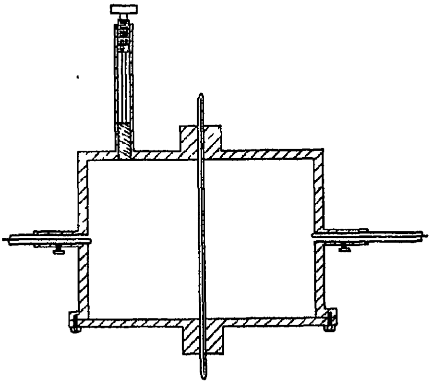


FIG. 1.—E_{010} Resonator used at 10 cm.

As an alternative a direct method of measuring the absorption coefficient (κ) was developed. The apparatus is shown in Fig. 2. The wave is fed from a coaxial line into a rectangular waveguide whose dimensions are such that when it is filled with water only the H_{10} mode can be transmitted. After passing through the water column the wave is again picked up by a probe and taken by a coaxial line to the mixer. The attenuation is measured by comparing by a substitution method the attenuation produced by various lengths of water column, with the attenuation of a piston attenuator. The receiver is only used as a monitor and does not have to be calibrated.

Consider a rectangular waveguide lying along the z axis of a Cartesian co-ordinate system, and which is bounded by the planes $x = 0$, $x = a_1$, $y = 0$, $y = b_1$, and which is filled with a material of propagation constant

$$p_1 = n_1 - \mathrm{j}\kappa_1.$$

The electric field of an H_{10} wave is represented by the equation

$$E_y = \frac{\mathrm{j}\omega a_1}{\pi} \cdot \sin\frac{\pi x}{a_1} \cdot e^{\mathrm{j}\omega(t - p_1 z/c)}$$

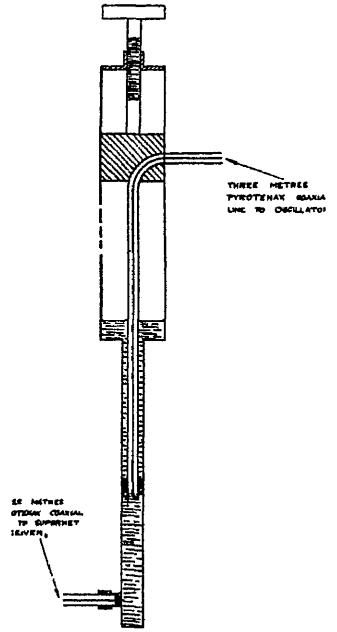


FIG. 2.—Wave guide attenuator at 10 cm.

The quantity which is measured is κ_1, the exponential absorption coefficient in the guide. The walls of the tube affect the propagation constant of the medium so that the relation $p = n - \mathrm{j}\kappa$ of free space is replaced by the relation

$$p_1^2 = (n_1 - \mathrm{j}\kappa_1)^2 = \epsilon_1 - \mathrm{j}\epsilon_2 - \frac{k_1^2 c^2}{\omega^2}$$

$$(k_1 = \pi/a_1 \text{ for the } H_{10} \text{ wave}).$$

Similarly in a second experiment with a guide of width a_2 one has

$$p_2^2 = (n_2 - \mathrm{j}\kappa_2)^2 = \epsilon_1 - \mathrm{j}\epsilon_2 - \frac{k_2 c^2}{\omega^2}$$

so that

$$\epsilon_1 = \frac{\frac{c^2}{\omega^2}(k_1^2\kappa_1^2 - k_2^2\kappa_2^2)}{\kappa_1^2 - \kappa_2^2} - (\kappa_1^2 + \kappa_2^2)$$

and

$$\epsilon_2 = 2\kappa_1\left[\epsilon_1 - \frac{k_1^2 c^2}{\omega^2} + \kappa_1^2\right]^{\frac{1}{2}}.$$

This method affords a satisfactory way of measuring both n and κ. As can be seen from the form of the equations the method is a practicable adaptation of the ideal method of determining κ by the attenuation in free space and n by measuring the critical wavelengths of propagation down a guide. The quantities which have to be measured accurately are the width of the guide and the wavelength. These are the two quantities which can be measured with the greatest precision, and the latitude in the choice of a_1 and a_2 is sufficient to allow the two measurements of the attenuation in the guide to be approximately independent measurements of n and κ, the optical constants in free space.

Measurements at 1·25 cm.—The technical difficulties of accurate measurement at 1·25 cm. are considerable. The oscillators are more sensitive to changes in load, temperature and electrode voltages. A very high degree of frequency stability must be maintained if one is to use the conventional I.F. amplifier of 2-3 mc. bandwidth.

After trying several other arrangements the apparatus of which a block diagram is shown in Fig. 3 was found to be satisfactory and capable of measuring κ to about 1 %.

The transmitting valve is frequency modulated by a saw-toothed voltage applied to the reflector. The wave is monitored by a directive feed section and passes through a window to the cell in which the water to be measured is maintained at a constant temperature, by water circulating from a thermostat.

Various designs of cell are shown in Fig. 4; details of its design will be discussed later. A section of flexible waveguide enables the thickness of the water through which the wave has to pass to be varied by means of the micrometer head. Having passed through the water and entered the wave guide through a second window the wave proceeds to a conventional mixer and I.F. amplifier.

The mean current in the cathode circuit of the second (diode) rectifier is measured on a high resistance microammeter after the standing diode current has been balanced out with a variable back bias. The pulse is also viewed on a C.R.O. whose time base is locked to the saw-toothed voltage originally used to modulate the frequency of the transmitting valve.

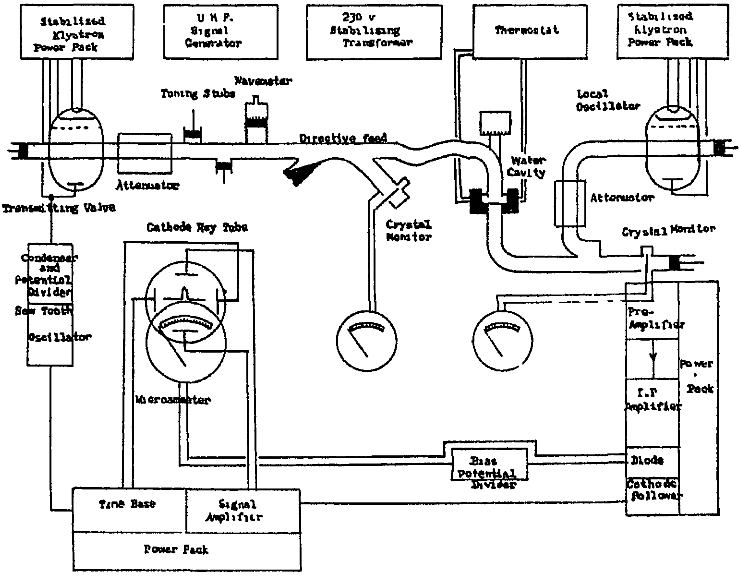


FIG. 3.—Block diagram of 1·25 cm. apparatus.

The pulse seen on the C.R.O. represents the frequency response curve of the I.F. amplifier, since amplitude modulation can be neglected if the mean reflector potential is suitably chosen. Frequency drift by either the transmitter or local oscillator, merely changes the phase of the pulse without changing its shape and is thus to a first approximation without effect on the mean diode current. In practice the pulse is kept at a particular phase by manual control whilst the diode current is read. In this way the difficulties of stability and calibration inherent in an absorption measurement are at least transferred from the 1·25 cm. transmitter to the 60 mc. I.F. amplifier where they are easier to deal with. The response curve of the amplifier is measured with a signal generator and found to be linear with voltage in the range used.

Design of the Cell.—In the first cell used (Fig. 4*a*) the wave issuing from the top window spreads into the wide circular pool of water until it reaches the conducting walls. During this section of its path its attenuation has a geometrical as well as a dielectric origin. Once it has reached the walls further spreading is prevented, and the wave becomes effectively a plane wave in a free dielectric. This effect could easily be observed; as the thickness of water in the cell was increased, the apparent absorption coefficient fell until it ultimately reached a steady value. As the necessary water thickness was rather large, a different design was finally adopted (Fig. 4*c*) in which the wave is always confined to a rectangular wave guide only slightly larger than the standard guide. In this way the non-divergent state is more quickly reached and the accuracy of measurement is considerably increased. Even with this cell difficulty was experienced at temperatures above 60° C. (where n is larger and κ smaller) in obtaining a sufficiently long water path over which the apparent value of κ did not change. In order to reduce the effect of standing waves the windows were made of marble (permittivity $\sim$ 8) cut so as to act as quarter wave matching transformers.

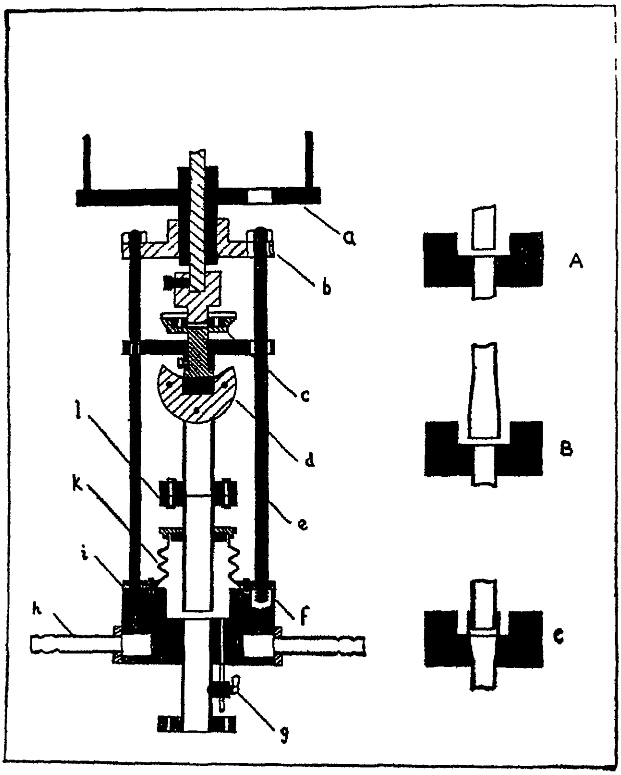


FIG. 4.—Water cell for use at 1·25 cm.

(*a*) Micrometer head; (*b*) micrometer rest; (*c*) ball race; (*d*) cut-away section of horizontal wave-guide with flange; (*e*) one of three supporting pillars; (*f*) copper block; (*g*) drain cock; (*h*) water supply from thermostat; (*i*) rubber gasket; (*k*) metal bellows; (*l*) detachable wave-guide head.

Results.

The results obtained are in reasonable agreement with the values given by von Hippel for 20° C. to 40° C.

$\lambda = 10{\cdot}00$ cm.

Temp. °C.	n (Resonator).	n (Waveguide).	κ (Waveguide).
0·0	8·99 ± 0·09	9·03 ± 0·05	1·37 ± 0·01
5·0	—	9·02	1·15
10·0	—	8·89	0·98
15·0	—	8·85	0·84 (5)
20·0	8·86	8·83	0·74
25·0	—	8·81	0·63
30·0	8·68	8·78	0·56
35·0	—	8·70	0·49
40·0	8·50	8·53	0·43 (5)
50·0	8·28	—	0·34 (5)
60·0	8·09	—	0·28
70·0	7·87	—	0·23
80·0	7·71	—	0·20
90·0	7·55	—	0·17
100·0	7·38	—	0·15

The variation of κ with temperature at $\lambda = 1{\cdot}26$ cm. is shown in Fig. 5. The estimated error is ± 1 % below 30° C. rising to ± 2½ % at 80° C.

The most comparable results are those of Saxton and Lane. There

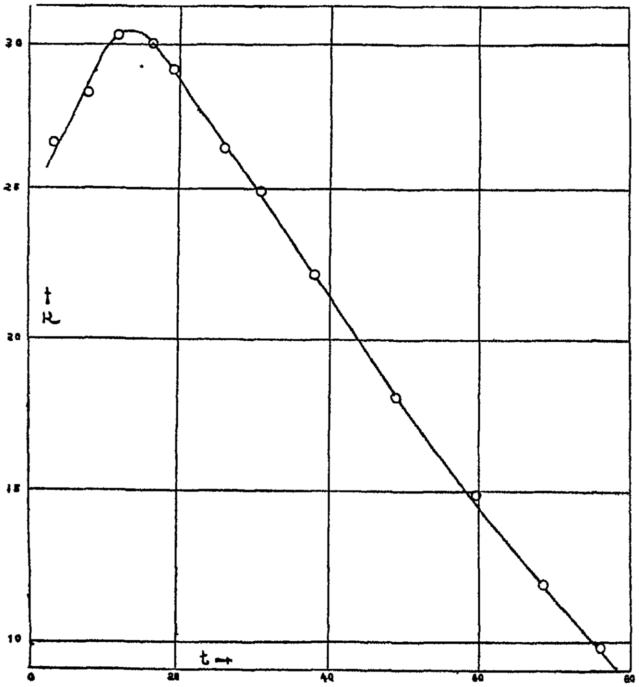


FIG. 5.—Curve of κ against temperature at 1·26 cm.

is excellent agreement as to the temperature at which κ goes through its maximum value and between the absolute values of κ above 25° C. Our results, however, show the maximum to be much sharper and considerably higher than the value found by Saxton and Lane. The reason for this discrepancy is not known. Our value for κ would be too high if a sufficiently large proportion of the wave energy were transmitted in a higher mode than the H_{10} mode which alone can be propagated in the empty guide. Assuming that half the energy were to appear in another mode, one can estimate the error in κ to be not more than 2 %. In any case, this explanation cannot account for the lower values found below 10° C.

The work described in this report has been carried out under a research contract for the Director of Scientific Research, Admiralty, and the authors wish to thank the Board of Admiralty for permission to publish.

The Clarendon Laboratory,
Oxford.

---

[1] Von Hippel, *O.S.R.D. Report*, 1197, Dec. 1942.

[2] Saxton (*in the course of publication*).

[3] Saxton and Lane (*in the course of publication*).

[4] Willis Jackson, *Ministry of Supply Report*, 287/Gen./35, Nov. 1941.

[5] For a general account of measurements up to 1938 see Muller, *Ergeb Exakt. Naturwiss.*, 1938, **17**, 164.

WAVE GUIDE MEASUREMENTS OF DIELECTRIC ABSORPTION OF SOLUTIONS OF POLAR SUBSTANCES IN NON-POLAR SOLVENTS.

By H. Warburton Hall, I. G. Halliday, W. A. Johnson and S. Walker.

Received 21*st March*, 1946.

Since the wave guide test components for centimetric radio engineering have tended to become more readily available, it was decided to examine the possibility of applying them, with modifications, to the measurement of the dielectric absorption of liquids.

Suitable electrical components were obtained and a preliminary attempt has been made to use them to investigate the variation of dielectric loss with viscosity at room temperature and at a frequency of approximately 10^{10} c.p.s. The liquids tested were dilute solutions of nitrobenzene and similar polar molecules in a two component mixture of non-polar solvents. The viscosity of the solution was varied by varying the proportions of the two solvents; one being a very viscous liquid.

Wave Guide Apparatus.

Broadly speaking there are two main groups of methods of measuring dielectric properties at the frequencies used in these experiments, *viz.*,

(*a*) Transmission line methods.
(*b*) Cavity resonator methods.

It was decided in the first instance to examine the possibility of using method (*a*) in conjunction with wave guides as the necessary components were readily available and adaptable. Further, as outlined below, the method contains no theoretical difficulties in obtaining absolute values. The rigid application of the method, however, involves, as described later, considerable computational difficulties and could only be justified for calibration purposes. For the present investigation these difficulties were considerably reduced by applying good approximations and with them the apparatus becomes suitable for routine measurements.

Briefly, the apparatus consists of a length of liquid filled guide terminated by a variable reactance; in practice, a sliding reflector. Fig. 1 illustrates the various sections of the line. These sections were connected together by hand screwed couplings.

The wave guide used was standard British copper guide with a cut-off wavelength of 5·08 cm.

The reflex Klystron, fed by a stabilised power pack was stable with regard to frequency and output once it had reached a steady temperature.

The attenuator provided variable attenuation sufficient to isolate the Klystron from the load on the remainder of the line.

The wave meter was an H_{10} cavity type.

The standing wave indicator was a special precision instrument developed at the Royal Aircraft Establishment and will be described elsewhere.

The output from the indicator crystal was taken directly to a moving coil galvanometer with a current sensitivity of about 10^{-2} microamps per mm. at 1 metre. By means of a vernier and scale on the indicator the position of the probe could be read to 1/20 mm.

The test cell consisted of a 6·1 cm. length of guide partitioned off from the rest of the line by means of thin mica windows fixed with glue. It was filled and emptied through two 1 mm. dia. holes in the centre of the long sides of the guide.

The sliding reflector, consisting of a choke plunger backed by a short, was moved, relative to the cell, within the guide, by a rack and pinion attachment. A singular value of the cell and termination impedance could be obtained by replacing the length of guide containing the reflector by a shorting plate. If the couplings were made hand tight replacements in the line gave repeatable results.

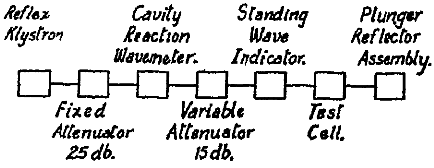

Reflex Klystron
Cavity Reaction Wavemeter.
Standing Wave Indicator.
Plunger Reflector Assembly.
Fixed Attenuator 25 db.
Variable Attenuator 15 db.
Test Cell.

FIG. 1.

Procedure for the Electrical Test.—The instrument was switched on for about one hour before the test in order that it might reach a steady temperature. During this time the frequency was checked with the wave meter and adjusted if necessary. With the cell empty, the value of the standing wave minimum was noted and unless there was some misalignment of the guide it was within ¼ mm. of the galvanometer zero, indicating no measurable loss.

Then with the cell filled (care being taken to eliminate any air bubbles by viewing the liquid through the mica windows) a reading was taken of the standing wave maximum and minimum value as indicated by the galvanometer deflections and of the position of the maximum on the standing wave indicator scale (by bracketing on the galvanometer scale).

From the galvanometer deflections the standing wave ratio was obtained whilst from the position of the maximum, knowing the wavelength (previously obtained by traversing the standing wave) of the radiation within the air-filled guide, the phase of the wave at the cell face was calculated.

By calculation from these two measurements, standing wave ratio and phase, for a number of positions of the reflector the input impedance of the test cell and its termination was found. The attenuation constant was then derived from the impedance figure and the phase constant from a measurement in which the shorting metal plate is used.

The electrical measurements were made at room temperature in a small inner laboratory where temperature fluctuations were comparatively low, ranging between 19° and 22° C.

Theory of the Electrical Measurements.

The main steps in the theoretical analysis of the electrical measurement are given in the attached appendix, but a brief description of the method of evaluation is as follows.

The components ϵ' and ϵ'' of the complex dielectric constant can be expressed in terms of the propagation constant γ for the liquid filled guide, λ the wavelength of the radiation in free space and λ_c the "cut off" wavelength of the particular wave guide size used, by the equation :

$$\gamma = \frac{2\pi}{\lambda}\sqrt{\left(\frac{\lambda}{\lambda_c}\right)^2 - \epsilon' + i\epsilon''} \quad . \quad . \quad . \quad . \quad \text{(i)}$$

and γ in terms of α the attenuation constant and β the phase constant for the liquid filled guide by :

$$\gamma = \alpha + i\beta. \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

For the guide used in these experiments $\lambda = 3{\cdot}273$ cm., $\lambda_c = 5{\cdot}080$ cm. and γ_0 (*i.e.* for air-filled guide) $= 1{\cdot}47_2 i$ by using equation (i).

In order to evaluate α the standing wave ratio (u) and corresponding phase (v) readings are converted into the real and imaginary parts of a complex impedance Z (*i.e.* $Z = r + ix$) by substituting them in the standard relations for a terminated line :

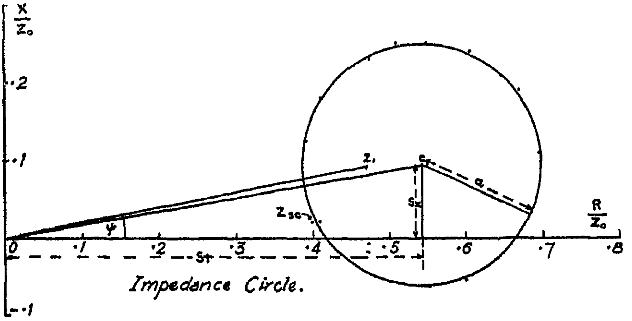


FIG. 2.

$$(r - c \cot h\, 2u)^2 + x^2 = c^2 \operatorname{cosec} h^2\, 2u$$
$$r^2 + (x + c \cot h\, 2v)^2 = c^2 \operatorname{cosec} h^2\, 2v$$

where c is a constant for the scale used in plotting. r = the resistive component of Z. x = reactive component of Z and from the solution of these equations the corresponding values of r and x are obtained; in practice this is done by using an impedance calculating chart. The series of values of r and x obtained from the measured values of (u) and (v) are plotted and a circle, centre C, as shown in Fig. 2 is fitted to the points.

The radius a of the circle and the real co-ordinate of point C_r are measured and their ratio substituted in the following equation :

$$\frac{S_r}{a} = \cosh 2\alpha l \cos^2 \psi \text{ (see appendix)} \quad . \quad . \quad \text{(iii)}$$

where $\cos\psi = \dfrac{S_r}{\sqrt{S_r^2 + S_x^2}}.$

Hence the value of α is obtained.

[illegible] β can be obtained from the singular value (Z_{SC}) of the impedance Z_A with the help of a graph of the function $\frac{\tan h\,\gamma l}{\gamma}$ previously computed for an appropriate range of values α and β since:

$$Z_{\mathrm{SC}} = \frac{\gamma_0}{\gamma} \tan h\,\gamma l \qquad . \quad . \quad . \quad . \quad \text{(iv)}$$

substitution of α and β in (i) then gives ϵ' and ϵ''.

Viscosity Measurements.—The coefficient of viscosity of each test sample was determined after the polar material had been added and as soon as possible after the electrical test. The measurements were made with a series of vertical capillary efflux type viscometers in a water bath maintained at a temperature of 25° C. The series were sufficiently large to cover the wide range of viscosities used and to provide overlapping for the comparison of one viscometer with the next in the series. Standard liquids such as distilled water, toluene, benzyl alcohol and sucrose solutions were used for calibration purposes. The viscometers were cleaned successively with benzene, ethyl ether or acetone and occasionally with chromic acid solutions and dried by means of cleaned compressed air. It was estimated that the error of measurement was within 1 %.

Selection of Substances for Test.—Since one of us (H. W. H.) had been interested for some time in the problem of obtaining dielectric absorption maxima, for simple solutions of liquid dielectrics, with variation of internal friction at constant temperature and frequency in place of varying frequency or temperature and had previously used polymerised styrene as a low loss material in dielectric measurements of mixed dielectrics, it was decided to apply the apparatus to this problem and to use purified styrene polymerised to various degrees to provide a variable viscosity solvent. In addition it was hoped to use the results to examine the Debye relation, and its modification involving Stokes' equation, for molecular relaxation times and that some information might be gained with regard to change in absorption with degree of polymerisation of the styrene itself.

However, initial experiments gave rather complex results. In addition the styrene itself gave a somewhat greater loss than was anticipated. It was decided therefore to use mixtures of strictly non-polar solvents in place of styrene for the initial series of tests and to return to styrene later.

The non-polar solvents selected were benzene, carbon disulphide, carbon tetrachloride, or hexane and medicinal paraffin.

The polar substances tested were nitrobenzene, α-nitronaphthalene and 4-nitro-diphenyl.

The bulk samples of the non-polar solvents were tested for dielectric loss and if necessary purified until they exhibited no measurable loss at the test frequency.

The results of tests of absorption of concentrations of nitrobenzene in benzene ranging from 0·5 to 6 g. per 100 c.c. of solution were plotted and a linear relation between attenuation factor and concentration was disclosed. From these results and from the fact that it gave a reasonably sized impedance circle a concentration of 4 g. of nitrobenzene per 100 c.c. of solution at room temperature was used throughout and the concentrations of the other polar substances were adjusted to give the same number of polar molecules per c.c. of solution.

Experimental Results.

Since the interest was principally that of detecting broad changes of dielectric absorption, that is in α, with viscosity and since the variation of β, the phase constant, was small throughout the range of tests made, the measured values of attenuation constant, α, were plotted against the viscosity of the solutions and the results are shown graphically in Figs. 3, 4 and 5.

The curves of Fig. 3 are for 4 g. of nitrobenzene to 100 c.c. of solution for all points. The terminal points for each curve are for a single solvent whilst all other points are for mixtures of two non-polar solvent solutions. It will be seen that three out of the four curves exhibit a maximum, that the value of the maximum and its position on the viscosity scale varies with each low viscosity solvent (*e.g.* hexane, carbon disulphide, benzene) used (that is there is a "solvent effect") and that as the value increases the maximum tends to occupy a position of lower viscosity. The fourth curve, that for carbon tetrachloride, suggests that under the conditions of test the viscosity of the single solvent solution is not low enough to provide conditions for maximum absorption. The fact that the viscosities for maximum absorption for the various curves of Fig. 3 are not identical suggests, in common with the reported results of other workers, that for solutions such as those examined here the coefficient of viscosity of flow for a liquid is only approximately a measure of the coefficient of internal friction (ξ) used by Debye to define molecular relaxation time τ, *viz.*, $\tau=\xi/2kT$ and that some modification involving the nature of the solvent molecule (*e.g.* polarisability) and its interaction with the polar molecule is necessary. By further work, now in progress, involving larger polar and non-polar molecules at a series of temperatures, it is hoped to gain further information on this approximation and other factors involved in molecular relaxation times.

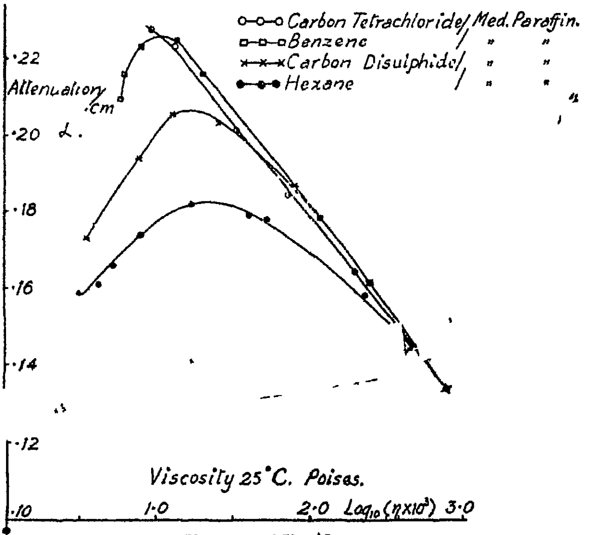


FIG. 3.—Nitrobenzene in mixed solvents, 4 gms./100 c.c. of solution.

In Fig. 4 curves are shown for nitrobenzene, α-nitronaphthalene, 4-nitrodiphenyl and a mixture of equal parts of nitrobenzene and nitronaphthalene in benzene/paraffin solutions. From these it will be seen that compared with nitrobenzene the larger nitro-naphthalene molecule, as would be expected, tends to give a maximum absorption at a lower viscosity, the shift on the log scale of the graph being from 1·05 to 0·90. Applying the Stokes' equation to the Debye relation for τ at the maximum a ratio in "hydrodynamical" radii for the two molecules of approximately 1 . 1 is obtained. In this figure a curve of ϵ'' against log viscosity is given to indicate the small change of β compared with that of α in these tests.

From the results shown in Fig. 3 it was noticed that although only four solvents were used it was just possible that the points for single solvents might lie on a smooth curve forming a boundary curve to those for the mixed solvents. In order to examine this possibility a number of other non-polar solvents were tested; nitrobenzene was again used as the polar solute, and the results are given in Fig. 5. This figure shows that with the exception of cyclohexane and possibly heptane the points do tend to lie on such a curve. The possibility of this result also being explained by the polarisabilities of the various solvents used is being examined.

It appears from the results of this investigation that frequencies of the order used should be particularly suitable for dielectric studies, at room temperature, of the smaller polar molecules.

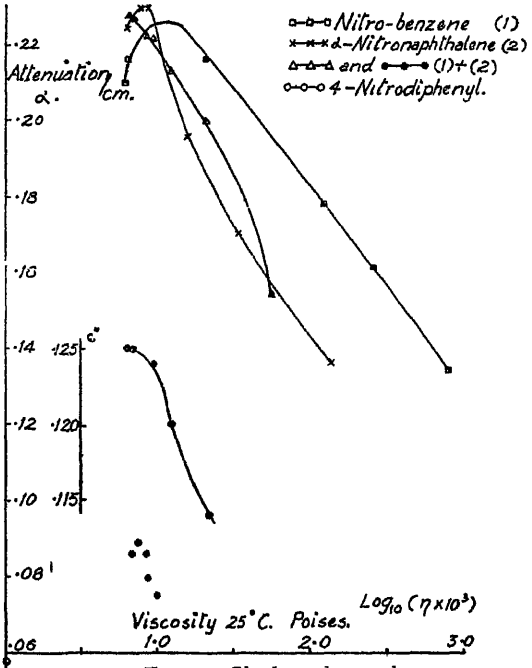


FIG. 4.—Single and mixed polar solutes in benzene/med. paraffin solutions.

Summary.

Preliminary results are reported of attempts to use wave guides to measure at room temperature and at a frequency of approximately 10^{10} c.p.s., the variation of dielectric absorption with viscosity of dilute solutions of polar molecules in two component mixtures of non-polar solvents. Standard radio engineering test components with modifications were used. The number of polar molecules per c.c. of solution was kept constant and the absorption measured at approximately constant temperature in solutions of different viscosities; the viscosity being varied by altering the proportion of high viscosity component in the mixture.

The polar substances examined were nitrobenzene, α-nitronaphthalene, and 4-nitrodiphenyl, the low viscosity solvents hexane, benzene, carbon tetrachloride and carbon disulphide and the high viscosity solvent medicinal paraffin. Only in the case of nitrobenzene was the full range of solvents used; the results with the other polar materials being limited by solubility difficulties.

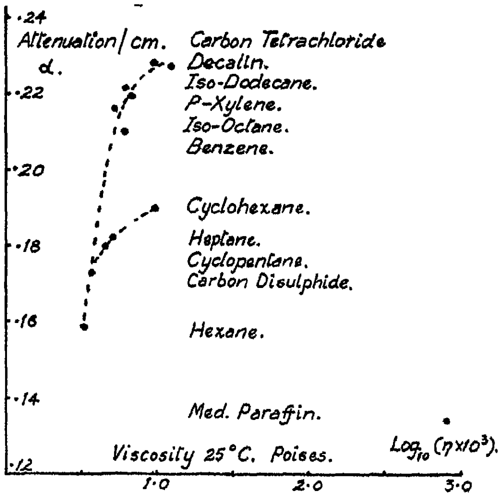


FIG. 5.—Nitrobenzene in various solvents.

In all cases where the paraffin was used with nitrobenzene the curve, at constant temperature, of absorption against log-viscosity tended to exhibit a maximum absorption but the position of this maximum on the viscosity scale and its value varied with the type of low viscosity component used with the paraffin.

Attempts to use purified styrene polymerised to various degrees to provide a series of viscosities from one material gave more complex results indicating the need for further investigation.

The test results suggest that before the Debye relation for relaxation time and internal friction incorporating Stokes' formula can be applied some modification involving the nature of solvent molecules used is necessary.

We are indebted to the Director General of Scientific Research, Ministry of Aircraft Production, for permission to publish these results; to Professor W. F. K. Wynne-Jones and Dr. R. J. W. Le Fevre for their continued interest and encouragement, to Dr. Mardles and Mrs. T. Savage for their help with the viscosity measurements, and to Professor E. E. Turner, F.R.S., for a gift of 4-nitrodiphenyl.

APPENDIX.

Impedance Circle Analysis.

A length of guide filled with a dielectric whose magnetic and electric parameters μ' and μ'', ϵ' and ϵ'' are independent of field strength is effectively a linear four terminal network. The input impedance of such a network is a linear function of its terminating impedance. By the properties of the bilinear transformation, if this terminating impedance is made to vary as a circle when plotted in the complex plane, then the input impedance will also vary as a circle. The most convenient variable impedance which meets such conditions is a reactance taking all values from $-\infty$ to $+\infty$, such as that provided by an air-filled guide of variable effective length.

We will assume in what follows that μ' is equal to unity and μ'' zero.

It would be possible to use a singular point on the circle, where the terminating impedance is either zero, or infinity. From the experimental point of view, however, the full circle is preferable since it provides a safeguard in that any one point measured with less precision than the others will be seen to be off the circle and may be ignored.

We have three characteristics of the input impedance circle: the radius, and the two co-ordinates of its centre. A fourth is provided by the length of the sample. It would then in theory be possible to obtain the two unknown parameters ϵ' and ϵ'' immediately from the equation for the circle.

The input impedance Z_A of a line of length l characteristic impedance Z_1, propagation constant γ and terminated by an impedance Z_2 is:

$$Z_A = Z_1 \frac{Z_2 + Z_1 \tanh \gamma l}{Z_2 \tanh \gamma l + Z_1} \quad . \quad . \quad . \quad . \quad (\text{v})$$

we set $\gamma = \alpha + i\beta$ where α = attenuation constant and β = phase constant.

In our case Z_2 varies reactively from $+\infty$ to $-\infty$ and is equal to $Z_0 i \tan \beta_0 l_0$ where the suffix " o " refers to the empty guide of the termination.

From (v) we may obtain relations for the radius and centre of the input impedance circle. They are

$$\text{Radius A} = \frac{R}{\sinh 2\alpha l \cos \psi + \sinh 2\alpha l \sin \psi} \quad . \quad . \quad . \quad (\text{vi})$$

$$\text{Real co-ordinate of centre B} = R \frac{\cosh 2\alpha l \cos^2 \psi + \cos 2\beta l \sin^2 \psi}{\sinh 2\alpha l \cos \psi + \sinh 2\alpha l \sin \psi}. \quad (\text{vii})$$

$$\text{Imaginary co-ordinate of centre C} = R \frac{(\cosh 2\alpha l - \cos 2\alpha l) \sin \psi \cos \psi}{\sinh 2\alpha l \cos \psi + \sinh 2\alpha l \sin \psi} \quad (\text{viii})$$

where $Z_1 = R e^{i\psi}$ in complex polar co-ordinates; and it is known that, for the H_{01} mode, $Z_1 \gamma = i\omega\mu$ = a constant.

It will be seen that to compute a large number of results from these relations would be tedious in the extreme. However as we were primarily interested in α we investigated the properties of the circle which were dependant on α to the first order.

The particular quantity to be avoided was the height of the circle above the axis, since it is always small, and greatly affected by any change in wavelength. In what follows thiis characteristic only occurs as a small correction :

From (vii) and (vi) we have :

$$B/A = \cosh 2\alpha l \cos^2 \psi + \cos 2\beta l \sin^2 \psi. \quad . \quad . \quad . \quad (ix)$$

For the circles we have obtained it was found that ψ is always fairly small, and therefore $\cos^2 \psi$ was nearly unity. It will be seen then that the real co-ordinate of the centre divided by the radius of the circle is a good measure of α.

For a typical circle B = 0·618, A = 0·173 and $\cos \psi = 0{\cdot}985$. B/A becomes 3·573. If we simply set

$$B/A = \cosh 2\alpha l \quad . \quad . \quad . \quad . \quad . \quad . \quad (x)$$

we find $2\alpha l = 1{\cdot}946$.

Now $\cos 2\beta l$ cannot be greater than unity. To find the approximate maximum error in (x) let us set $\cos 2\beta l$ equal to unity. With $\cos^2 \psi = 0{\cdot}97$, $\sin^2 \psi = 0{\cdot}03$

$$\cosh 2\alpha l = 3{\cdot}640 \quad \text{and} \quad 2\alpha l = 1{\cdot}965.$$

$2\alpha l$ is then in error by 1 %. Since the angle ψ did not vary in all our circles by more than 10 % this error may be reduced to less than 0·1 % by dividing B/A by $\cos^2 \psi$. ψ may be taken for the purpose of this correction to be the angle between the line joining the circle centre to the origin and the real axis, since Z_1 is always close to the centre.

The experimental error in computation and measurement is difficult to assess, due to the complicated method of computing the impedances, and the averaging out process inherent in fitting a circle through the points. All the lengths measured by the standing wave indicator were correct to 0·1 %. Any error in the wavelength would be included in the variation of ψ. The known accuracy of the polar impedance chart used was 0·5 %. The error in measurement of B and A from the circle would be within 0·5 %. The total estimated error in α then will be less than 1·5 %.

The computation of β from the circle depends critically on the imaginary co-ordinate of the circle centre. In these experiments this co-ordinate was always small making the exact determination of β directly from the circle difficult. Therefore for a number of tests the method of obtaining a singular value (Z_{sc}) of Z_A was employed in which the expression, $Z_{sc} = \frac{\gamma_0}{\gamma} \tanh \gamma l$ and a chart of the function is used. Since ϵ'' is directly proportional to $\alpha\beta$ from (1) and in a typical curve, (Fig. 4) the maximum variation in β was not greater than 0·2 times that of α, then the change in α may be regarded as the principal measure of the change of ϵ''.

Materials Department,
Royal Aircraft Establishment,
S. Farnborough, Hants.

[1] Slater, *Microwave Transmission*, McGraw Hill Book Co.
[2] Willis Jackson, *High Frequency Transmission Lines*, Methuen.
[3] Lamont, *Wave Guides*, Methuen.
[4] Titchmarsh, *Theory of Functions*, O.U.P.
[5] Debye, *Trans. Farad. Soc.*, 1934, **30**, 679.

EXTENSION OF THE MEASUREMENTS OF DISPERSION AND ABSORPTION BY LIQUIDS, TO THE REGION OF CENTIMETRIC RADIO-ELECTRIC WAVES.

By P. Abadie.

Received in French on 22nd February, 1946, and translated by Dr. M. Welch.

Several years ago, we undertook measurements on dispersion and absorption on polar liquids in the Hertzian region, by varying the wave length λ and keeping the temperature constant. In this way it has been possible to obtain large segments of a curve for certain bodies, but at the shortest λ used (about 18 cm.), conditions were still rather far from the optical dielectric constant ϵ_0. Having been able to set up magnetrons covering a range of wavelengths of which the lower limit was about 2·5 cm. shortly before the war, we have made use of them particularly to complete the curves already obtained.

Method used and its Working.—Measurements had already been carried out for λ's between 10 and 2·8 cm. by an optical method[1]. After several attempts we gave up the use of this latter, as it required too much liquid, and also because of the large dimensions of the cells, we were far from realising the conditions necessary for obtaining correct results. We then applied the first method of Drude. This latter had been used by many experimenters, and in particular, measurements had been carried out using decimetre waves,[2] to λ = about 13 cm. and also centimetre waves,[3] to about 6·5 cm.

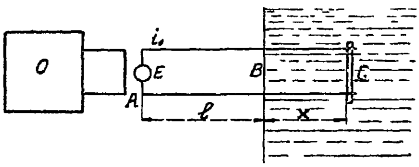


FIG. 1.

In our measurements we used a double line consisting of two coaxial cylinders. Before giving details of the apparatus, we recall the method of applying the experimental results to the deduction of the value of the complex dielectric constant $\epsilon' - j\epsilon''$ of the liquid studied. Let us consider, then (Fig. 1), a double line short-circuited at the end A, the first part consisting of a length l in air, the second part consisting of a variable length x in the liquid being studied, and finally being short-circuited at the end C which is situated in the liquid. The end A is coupled to an oscillator O, which we assume to act as an e.m.f. E, localised at A. Moreover we assume that the coupling is sufficiently loose to enable E to remain constant when x varies. Applying the classical equations for propagation along lines, and neglecting the resistance of the conductors we put

$$\sqrt{\epsilon' - j\epsilon''} = n(1 - j\kappa)$$

and $\lambda' = \lambda/n$ where λ is the wavelength in air. Assuming $l = k\,\lambda/2$ where k is whole, and designating the current in A by i_0, it can be seen that the deviations δ of an apparatus passing i_0 are given, if δ is proportional to $i_0{}^2$, by

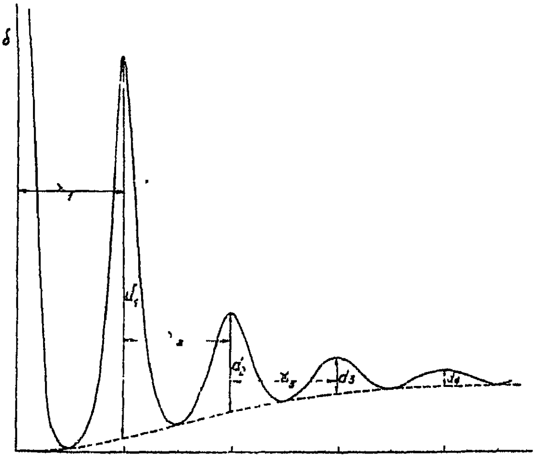


FIG. 2.

by
$$\delta = K\,\frac{1 + \tanh^2 2\pi\kappa\,x/\lambda' \,.\, \tan^2 2\pi\,x/\lambda'}{\tanh^2 2\pi\kappa\,x/\lambda' + \tan' 2\pi\,x/\lambda'} \quad . \quad . \quad . \quad (1)$$

---

[1] Baz, *Physik. Z.*, 1939, 394.

[2] von Ardenne, Gross and Otterbein, *ibid.*, 1936, **37**, 533-544.

[3] Knerr, *Physical Rev.*, 1937, **52**, 1054.

The curve $\delta = f(x)$ has the shape given in Fig. 2. To the first approximation, δ is a maximum for $x = \lambda/2, \lambda', 3\gamma'/2, \ldots$ and

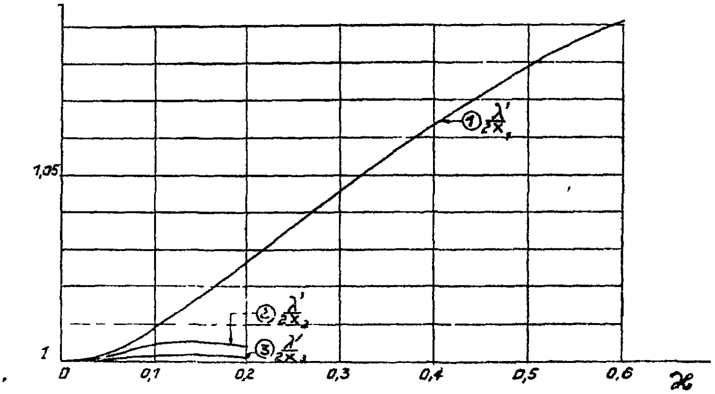


FIG. 3.

the quantities $d_1, d_2, \ldots$ are such that starting from a certain order k, we can put

$$\frac{d_k}{d_{k+1}} = \frac{d_{k+1}}{d_{k+2}} = \ldots = e^{2\pi\kappa}$$

which enables us to determine

$$n = \lambda/\lambda^1 \text{ and } \kappa.$$

In fact if δ is derived with reference to x, the equation

$$\tan 4\pi\, x/\lambda' + \kappa \tanh 4\pi\kappa\, x/\lambda' = 0$$

is obtained, which can be solved by successive approximations. The curves can then be traced from Fig. 3 in which x_1 is the abscissa of the 1st maximum, x_2 the distance between the 1st and 2nd maximum, . . . etc. Curve 1 has been used when the liquid is very absorbant, as for example in the case indicated in Fig. 4, for only the 1st maximum is known with much precision.

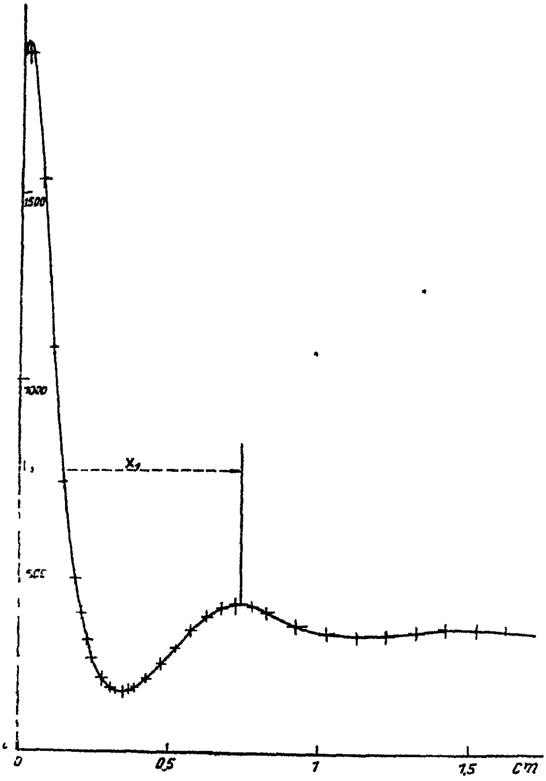


FIG. 4.

On the other hand, the quantities $d_1, d_2, \ldots$ of Fig. 2 decrease very quickly as soon as the absorption is considerable. From the point of view of experimental precision, conditions are bad if the values only starting from the order k are considered as above-mentioned. It is obviously better to use the first maximum, it being only necessary to introduce a correction factor. For this purpose we have drawn the curves of Fig. 5 taking $d_k = \frac{1}{\tanh^2 k\pi\kappa} - \tanh^2 k\pi\kappa$, a value very close to the real value.

So far we have assumed that the resistance of the line was negligible, We sought to correct for this by inserting in the scheme of Fig. 1 a resistance of value suitable to A. The preceding results are not sensibly modified,

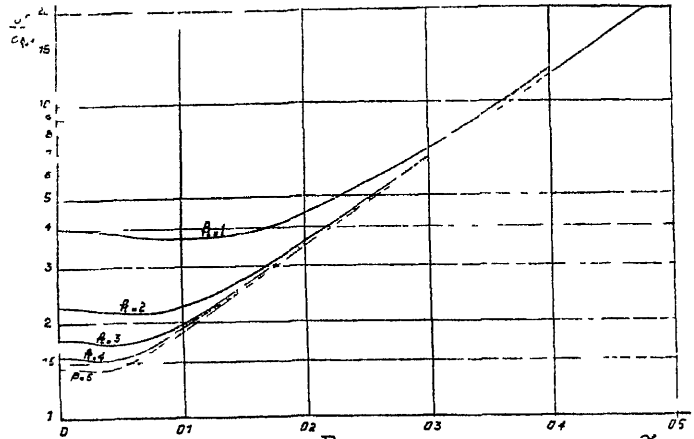


FIG. 5.

We have finally assumed that l (Fig. 1) was equal to a whole number of $\lambda/2$, in other words the air-liquid interface occurs exactly at a potential node. In order to see how the results are modified when this is not the case we have repeated the calculation assuming l has any value whatsoever, the line resistance being neglected. The rather complicated expression obtained is not readily susceptible to mathematical discussion, but we have traced three curves corresponding to $\kappa = 0{\cdot}15$, to $\tan 2\pi\, l/\lambda = 0$, and to $\tan 2\pi\, l/\lambda = \pm 0{\cdot}3$. The differences between the results given by these three curves are clearly less than those due to experimental error. It should be noted that the cases, $\tan 2\pi\, l/\lambda = \pm 0{\cdot}3$ are extreme cases from which we are always very far in practice. It was interesting to establish this point, for if one can refer to a case like that in Fig. 4, where the value of x_1 only, can be determined, it is much better to place the potential peak slightly to the interior of the liquid so as to be quite sure of its position.

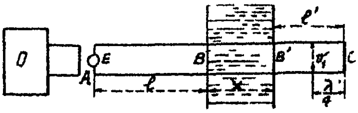


FIG. 6.

It can be seen from Fig. 4 that if the absorption is exceptionally strong, κ can only be obtained from the ratio d_k/d_{k+1}. We have then used the following method, which is identical with the optical method: After a part AB in the air (Fig. 6), the line allows a part BB′ of length x to remain in the liquid, and finally a part B′C in the air; C is a short circuit. The wave passing into the liquid in the path BB′, is somewhat weakened; a part penetrates into the region B′C, another part is reflected and returns towards B, there undergoing a new reflection, and comes back towards B′. If for a given κ, x is assumed to be so large that after this passage, the reflected path is negligible compared with that which has penetrated directly in B′C, and assuming that E is the e.m.f. induced in A, the excitation of the element B′C is proportional to $E \,.\, e^{-2\pi n\kappa x/\lambda}$. l' having a suitable value, the potential v is registered at a point distant $\lambda/4$ from C by means of a detector giving deviations δ proportional to the squares of potential applied. It was known that $\delta = Ke^{-4\pi n\kappa x/\lambda}$ and the law $\log \delta = f(x)$ had for an angular coefficient $-\ 4\pi n\ x/\lambda$, from which $n\kappa$ and consequently κ are obtained.

Apparatus.—Fig. 7 gives the outline of the arrangement used. The line is coupled by a capacity to the magnetron oscillator. The section of the coaxial line Tt' which can slide on the real line of measurement, suitably establishes the potential node; "r" is a fibre ring. D is a crystal detector moveable along the tube, for deriving the curves of Fig. 2. So far we have supposed that the apparatus gives deviations proportional to i_0^2; in reality the line is open at the end coupled to the oscillator, but, except the constant, nothing is changed in the calculations if one assumes that D gives deviations proportional to the square of the potential at the potential node on the line. The piston P, which assures the short-circuit of Fig. 1 is controlled by the micrometer screw V, and moves within the liquid. Fig. 7 represents the arrangement of the apparatus when a measurement of κ is being made by variation of the thickness of liquid traversed. D is not used. D′ is a crystal detector that can be made to register the displacements of P (short-circuit of Fig. 4) and is not used where curves such as that of Fig. 2 are recorded.

For our measurements, we used four complete installations of this pattern, functioning respectively in the regions of 3, 6, 11 and 18 cm. This last wavelength was produced by a tube with a positive grid.

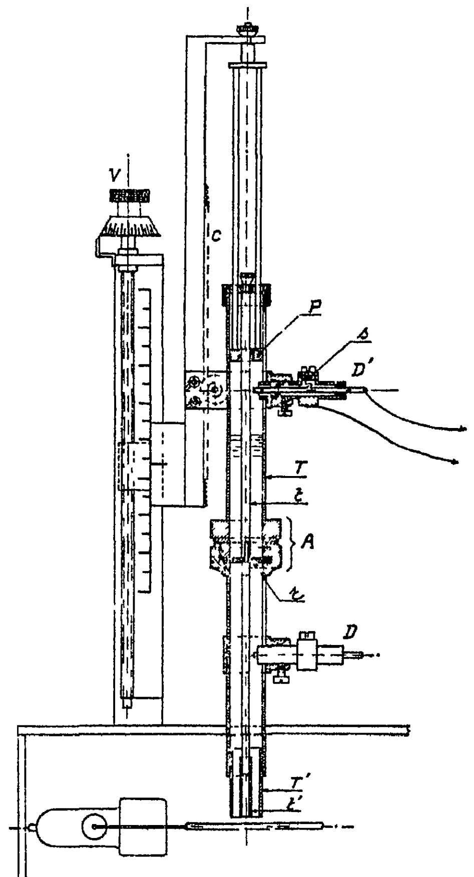


FIG. 7.

Some Experimental Results.

(1) **Benzene.**—For λ in the neighbourhood of 3, 6, 11 and 17 cm., we found $\epsilon' = 2{\cdot}29$ at least within 1 %; ϵ'' is practically zero.

(2) **Acetone** (propanone).—The liquid subjected to test (as well as nitrobenzene and propyl alcohol which will be dealt with later) had been purified by Laboratoires des Industries Chimiques de l'Etat. The results of the measurements have been given in Fig. 8.

The curves

$$\epsilon' = \epsilon_0 + \frac{\epsilon_1 - \epsilon_0}{1 + A^2\omega^2},\ \epsilon'' = (\epsilon_1 - \epsilon_0)\frac{A\omega}{1 + A^2\omega^2} \quad . \quad . \quad (2)$$

have also been traced, taking

$$\epsilon_1 = 21{\cdot}5$$

(determined experimentally), $\epsilon_0 = 2{\cdot}5$ (taken arbitrarily) and giving A a value such that the experimental values of ϵ' and ϵ'' fall as well as possible on the calculated curves.

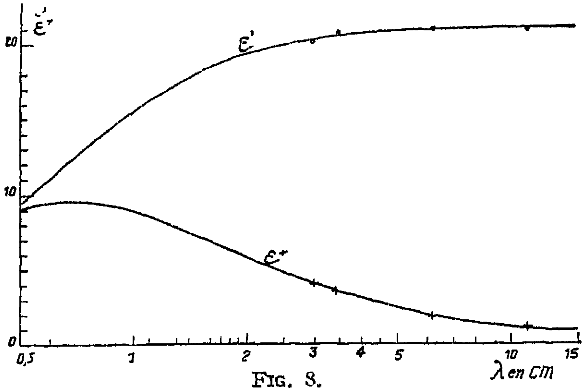


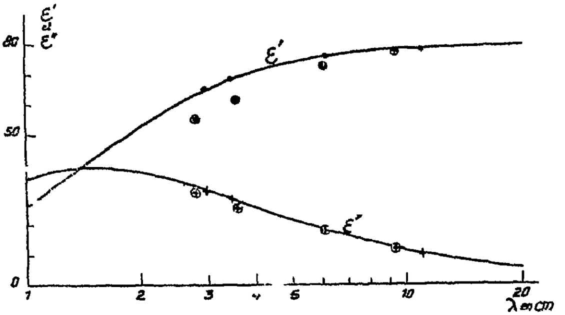
FIG. 8.

By extrapolation the wavelength λ_s corresponding to

$$A\omega = 1$$

(for which and ϵ'' is maximum) is deduced

$$\lambda_s = 0{\cdot}66 \text{ cm.}$$



⊛ + } Expt. points obtained by us for {ε', ε''} and

⊙ Ɔ } Points obtained by G. Bäz.

FIG. 9.

(3) **Water.**—The results are given in Fig. 9. The curves represented by the unbroken line have been calculated by taking

$$\epsilon_1 = 81, \quad \epsilon_0 = 2{\cdot}5,$$

and by choosing A as above. *By extrapolation* it is found that $\lambda_s = 1{\cdot}5$ cm.

The values obtained by Baz,[1] given in Fig. 9 are systematically different from ours for ϵ' and give the value $\lambda_s = 1{\cdot}85$ cm.

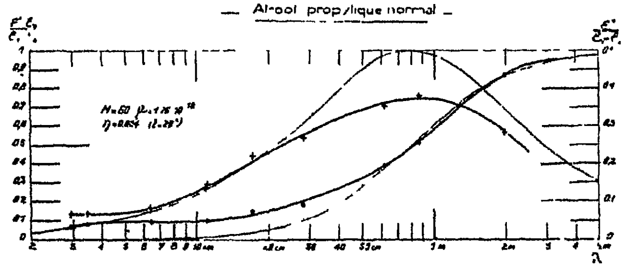


FIG. 10.

(4) **Nitrobenzene.**—The results, as well as those obtained with longer waves, are shown in Fig. 3 from the paper by Girard and Abadie entitled

" Experimental Curves of losses and of Anomalous Dispersion as a basis of a spectral method."

The experimental curves are shown in unbroken heavy line. The measurement on the shortest waves were delicate as we were in the region in which ϵ''/ϵ' was maximum. In order to draw the theoretical curves (represented by a fine line), the value of ϵ' is taken as 35·5 (determined experimentally), $\epsilon_0 = 2\cdot5$ and A is chosen as above. The experimental curve shows a distinct deviation from the theoretical curve.

(5) **Normal Propyl Alcohol.**—The results are given in Fig. 10. Here again the experimental curves are in heavy unbroken line; ϵ_1 has been taken $= 19\cdot9$ (determined experimentally) and $\epsilon_0 = 2\cdot2$. The results are given for the longer waves which led us to believe that a value very close to that of ϵ_0 might be obtained from the short waves which we had used. It can be seen, however, that this is not so; after a level stretch of ϵ', the curve starts to fall; moreover ϵ'' ceases to decrease for the shortest waves used. There is a very large deviation from the theoretical curves, which are traced by a fine line, the position of which we have arbitrarily fixed with respect to the λ axis. Again results analogous to octyl alcohol are found (see Fig. 1 of the note mentioned under " nitrobenzene ").

Conclusion.

The interpretation of certain of the preceding results is given elsewhere. We will give here only the following conclusion. For the liquids we have studied and for which we have traced a great part of the dispersion and absorption curves, it can be stated that their form deviates considerably from that of the theoretical curves calculated from equations (2). It is therefore dangerous to extrapolate the experimental results corresponding to only a fraction of the curve; the same applies even when this latter seems to coincide with part of the theoretical curve.

Laboratoire Nationale de Radioelectricité,
196 *rue de Paris,*
Bagneux (*Seine*),
France.

SOME MEASUREMENTS ON THE ABSORPTION OF CENTIMETRIC WAVES BY LIQUID DIELECTRICS

By F. J. Cripwell and G. B. B. M. Sutherland

Received 16*th May*, 1946

The study of electromagnetic radiation in the infra-red has yielded an immense amount of information about *intra*-molecular forces but comparatively little concerning *inter*-molecular forces, as the characteristic frequencies controlled by the latter lie in that intermediate region of the spectrum which includes the very long infra-red and the ultra-short radio waves. As a first step in the investigation of this region of the spectrum, it seemed best to see how far some of the techniques developed in the ultra-short wave radio field might be applied to problems involving intermolecular forces. In this paper we give a brief account of a method of measuring the absorption of centimetric waves using a wave-guide technique originally developed by Roberts and von Hippel [1] for 6 cm. waves.

---

[1] Roberts and von Hippel, *Physic. Rev.*, 1940, 57, 1056 (A).

Apparatus.—The waves are generated by reflector type Klystron oscillators and with the input designed to transmit only the H_{01} mode through a horizontal rectangular wave-guide, in which provision is made for variable attenuation, frequency measurement, power monitoring and detection of standing waves (Fig. 1). The far end of the wave-guide is bent to the vertical position and contains the liquid being investigated. The wave-lengths employed were 1·25 and 3·25 cm.

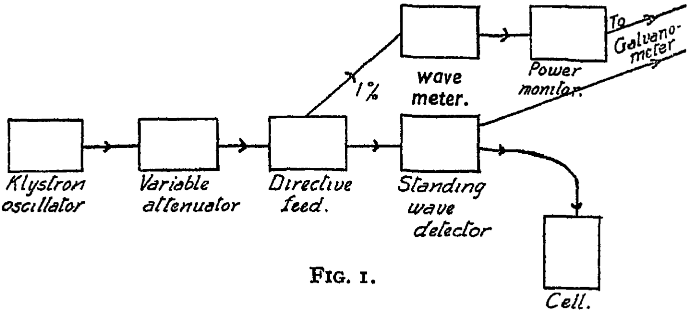


FIG. 1.

Theory.

Considerations of the input and output impedance of the cell show that for such a system—

$$\frac{\operatorname{Tanh}\gamma_2 d}{\gamma_2 d} = \frac{\left|\dfrac{E_{\min.}}{E_{\max.}}\right| - j\tan\dfrac{2\pi x_0}{\lambda_g}}{j\dfrac{2\pi d}{\lambda_g}\left\{1 - j\left|\dfrac{E_{\min.}}{E_{\max.}}\right|\tan\dfrac{2\pi x_0}{\lambda_g}\right\}} \quad . \quad . \quad (1)$$

Where γ_2 is the propagation constant of the liquid, d is the depth of the liquid $|E_{\min.}/E_{\max.}|$ is the standing wave ratio, x_0 is the distance of the first minimum from the surface of the liquid, and λ_g is the wave-length in the guide.

All the quantities on the right hand side are measurable and so γ_2 should be determinable. Unfortunately, the above equation cannot be solved directly. Roberts and von Hippel have used graphical methods but we have found it preferable to use a method of successive approximation. Thus if we put $\gamma_2 = a + jb$ or $\gamma_2 d = a' + jb'$ and let

$$\frac{\operatorname{Tanh}(a' + jb')}{a' + jb'} = x + jy$$

then solving for real and imaginary parts we get—

$$x = \frac{a'\sinh 2a' + b'\sin 2b'}{(a'^2 + b'^2)(\cosh 2a' + \cos 2b')}$$

$$y = \frac{a'\sin 2b' - b'\sinh 2a'}{(a'^2 + b'^2)(\cosh 2a' + \cos 2b')} \quad . \quad . \quad . \quad . \quad (2)$$

It turns out that rough values for a' and b' can readily be deduced and by successive trial and error it is then possible to determine a' and b' to the desired accuracy.

From γ_2, the propagation constant, the loss factor tan δ for the dielectric is then found by the following relations between ϵ' and ϵ'' the components of the complex dielectric constant $\epsilon_c = \epsilon' - j\epsilon''$ and a and b the components of $\gamma_2 = a + jb$:—

$$\epsilon' = \lambda^2/\lambda^2_c + \lambda^2/4\pi^2\,(b^2 - a^2)\ ;\ \epsilon'' = ab\lambda^2/2\pi^2$$

while $$\tan\delta = \epsilon''/\epsilon' \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Here λ is the wave-length in free space, and λ_c is the critical wave-length in the guide.

Results.

We have measured the tan δ values for a number of dilute solutions of polar liquids in benzene as these may be expected to conform most closely to the Debye theory. The results are given in the first three columns of Table I. The concentrations of the solutions used were between 0·01 and 0·1 moles per 100 cc. (except for toluene where the pure liquid was employed) and the results have been reduced to a theoretical concentration of 1 mole per 100 cc.

Now Debye's theory gives the following relation between tan δ and f, the frequency at which it is measured

$$\tan \delta = \frac{(\epsilon' + 2)^2}{\epsilon'} \cdot \frac{2\pi f \tau}{1 + 4\pi^2 f^2 \tau^2} \cdot \frac{4\pi \mu^2 N_c}{27kT} \qquad (4)$$

where τ is the time of relaxation of the solute molecule under the applied field, μ is the dipole moment of the solute molecule, ϵ' refers to the solution, N is Avogadro's number, c is the concentration in moles per cc., and k is Boltzmann's constant.

TABLE I.—(Temp. 18·5 ± 1° C.).

Solute in Benzene.	Value of tan. δ.		Relaxation Time $\tau \times 10^{12}$ sec.	Dipole Moment $\mu \times 10^{18}$ e.s.u.	
	1·25 cm.	3·25 cm.		Calculated.	Literature.
Benzophenone	1·43	2·25	18·1	2·81	2·98
Ethyl benzoate	0·77	1·15	17·6	2·02	1·82
Chlorobenzene	$0{\cdot}56_3$	0·561	$10{\cdot}8_5$	1·51	1·52-1·69
Methyl benzoate	0·781	0·919	13·0	1·83	1·84
Nitrobenzene	3·47	4·14	13·0	3·88	3·98
Cyclohexanone	2·20	2·15	10·5	2·90	2·8
Toluene	0·0295	0·0218	7·58	0·322	0·35-0·422
Nitromethane	1·97	1·02	4·36	2·81	2·74-2·85
Methyl acetate	0·602	0·291	3·73	1·58	1·67
Acetone	1·68	0·778	$3{\cdot}3_4$	2·6	2·74
Acetonitrile	2·13	0·919	2·5	3·4	3·2

Thus if ϵ_c is determined for a series of values of f and only one time of relaxation τ is involved, a curve of tan δ against f corresponding to (4) can be constructed to fit the results and so μ and τ can be determined. This curve will obviously have a maximum at $f = 1/2\pi\tau$. When only two values are available, Whiffen and Thompson[2] have suggested that a first approximation of τ can be obtained by considering the variation with τ of

$$\frac{\tan \delta_1}{\tan \delta_2} = \frac{\theta_1}{\theta_2}$$

where

$$\theta = \frac{2\pi f \tau}{1 + 4\pi^2 f^2 \tau^2}$$

Using this method we have calculated the values of τ and μ given in columns 4 and 5 of Table I. The agreement with the literature values for the dipole moments is seen to be reasonably satisfactory, and on the whole somewhat better than that obtained by Whiffen and Thompson.[3] The only literature values for times of relaxation are those of Holzmuller,[4] where at 24° C. he gives $16 \cdot 1 \times 10^{-12}$ for benzophenone, $6 \cdot 78 \times 10^{-12}$ for cyclohexanone and $2 \cdot 47 \times 10^{-12}$ sec. for acetone. The agreement is satisfactory, except for cyclohexanone.

---

[2] Whiffen and Thompson, *Trans. Faraday Soc.*, 1946, 42 A, 114

[3] We are indebted to these authors for communicating their results to us prior to publication. The particular compounds chosen were ones which had already been studied by Whiffen and Thompson (*Confidential Report*, May 1945) in order to compare the accuracy of our method with theirs.

[4] Holzmuller, *Physik. Z.*, 1937, 38, 574.

The estimated error in these results is 3-5 %, as judged by reproducibility at various depths. It should be remembered, however, that this method will be more inaccurate when low losses are being measured. The neglect of any loss due to solvent may tend to make our values of tan δ slightly high. This method has, however, the distinct advantage over the wave-guide method of Whiffen and Thompson[2] that no error arises because of ignorance of reflection coefficients at interfaces.

This work forms part of a programme of fundamental research sponsored by the Hydrocarbon Research Group of the Institute of Petroleum. We should also like to express our indebtedness to the Telecommunications Research Establishment, Malvern, for most of the apparatus, and to Professor Willis Jackson and Mr. Powles for many valuable discussions.

Cavendish Laboratory and Pembroke College,
Cambridge.

NOTE ON THE COMPARISON OF ENTROPY AND ENERGY CHANGES IN DIFFUSION AND DIELECTRIC RELAXATION PROCESSES.

BY D. L. LEVI.

Received 9th September, 1946.

In a recent paper by Whiffen and Thompson[1] the question was raised of the physical interpretation of negative entropy changes which had been found to characterise certain dielectric relaxation processes. Since the author has already found similar negative entropy changes to occur in diffusion processes in solids and has developed an interpretation of them in molecular terms, it seems worth while to extend this interpretation, if possible, to dielectric relaxation.

The values of the activation energy (ΔE) and entropy (ΔS) for dielectric relaxation are calculated from determinations of the variation of relaxation time (τ) with temperature. The calculation is based upon Eyring's equation :[2]

$$1/\tau = (kT/h) \,.\, e^{\Delta S/R} \,.\, e^{-\Delta E/RT}.$$

The corresponding energy and entropy changes for diffusion processes are based upon the relation :[3]

$$D = e\lambda^2 (kT/h) e^{\Delta S/R} \,.\, e^{-\Delta E/RT}$$

where D is the diffusion constant and λ is the mean free path of the diffusing molecule in the solid.

Tables I and II contain various values of ΔE and ΔS for the two kinds of processes. The data of Table I are taken from various sources ; those of Table II are based upon unpublished results of the author. The values for ΔS in Table II have been calculated by assuming $\lambda = 10 \times 10^{-8}$ cm. ; but the uncertainty involved in this assumption should not affect the ΔS values by more than about 2 entropy units.

The values of ΔE and ΔS for water vapour diffusion through solids can all be given a consistent physical interpretation in terms of the molecular configurations concerned in the transition state which precedes movement of the water molecule from one equilibrium position to the next. If formation of the transition state involves an increase in the number of degrees of freedom of the system, ΔS will be positive, and vice versa. Two causes of an increase in the number of degrees of freedom can be distinguished.

---

[1] Whiffen and Thompson, *Trans. Faraday Soc.*, 1946, **42** A, 114.

[2] Glasstone, Laidler and Eyring, *Theory of Rate Processes*, Chap. 9.

(*a*) If the lattice molecules (or segments) between the two positions of the water molecule require to be displaced in order to permit transfer, this will involve a disordering of the lattice, and ΔS will be positive. This process in diffusion is generally known as hole-formation. Since the displacement of lattice molecules in this way will require an activation energy, ΔE will on this account tend to be large, so that positive values of ΔS and large values of ΔE will occur together.

TABLE I.—ACTIVATION ENERGY AND ENTROPY FOR DIELECTRIC RELAXATION

Material.	ΔE (kcal./mole.).	ΔS (cal /deg./mole.).
Phenol-formaldehyde polymer [3]	84	208
Polyvinyl chloride (unplasticised) [3]	120	300
„ „ + 10 % tricresyl phosphate [3]	88	199
„ „ + 30 % „ „ [3]	66	156
Ethylene glycol phthalate [3]	53	112
Rubber [3]	31	62
Dioctyl phthalate in polythene [4]	15	33
„ „ polystyrene [4]	5	−2
Chloroform (in heptane solution) [1]	1	−3
o-Xylene [1]	1	−4

TABLE II.—ACTIVATION ENERGY AND ENTROPY FOR WATER VAPOUR DIFFUSION THROUGH SOLIDS.

Material.	ΔE (kcal./mole.).	ΔS (cal./deg./mole.).
Polyvinyl chloride +30 % tricresyl phosphate	48	120
Polythene	—	about 100
Linseed oil	13	11
A baking varnish	12	6
Perspex	10	−2
Cellulose acetate	8	−9

(*b*) In the case of solids which, for chemical reasons, cannot form hydrogen-bonds to water, the water is held by a configuration of the lattice that involves a considerable amount of ordering of the lattice, as is shown by the large, *negative* entropies of sorption. This applies to hydrocarbons (*e.g.* polythene, polystyrene) and to polyvinyl chloride. When diffusion occurs this ordered arrangement which holds the sorbed molecule must become disorganised. ΔS will then be large and positive, as also ΔE will be. This accounts for the high ΔS values of polyvinyl chloride and polythene given in Table II.

Negative values of ΔS (for which the transition state must involve a decrease in the number of degrees of freedom) are explained by assuming that the water molecule must be oriented in a particular direction before transfer takes place. This would reduce the number of degrees of freedom of the transition complex as compared with the equilibrium state of the molecule and so make ΔS negative. In these cases it will still be true that in certain phases of the *lattice* vibrations, the conditions will be especially favourable for transfer (this being essentially the hole-formation condition), but the present view implies that this favourable condition is attained more frequently by means of an oriented vibration of the water molecule, than by a suitable displacement of the lattice. It is noticeable that as ΔS for diffusion decreases, ΔE also decreases. This will be because as ΔE decreases, the water molecule can itself attain the requisite energy more frequently, so that it can take advantage of a larger number of lattice arrangements for the purpose of diffusing; thus the limiting factor then becomes not the formation of a suitable lattice arrangement, but a favourable direction of vibration on the part of the water molecule. This allows it to diffuse without the aid of any abnormal local disorganisation of the lattice. At the same time, the mere absence of any energy requirement for disorganising the lattice will keep ΔE low; so that we should always expect negative values of ΔS to be accompanied by low values of ΔE. The data of Table II bear out the foregoing interpretation of the diffusion process in detail.

---

[3] Results taken from Kauzmann, *Rev. Mod. Physics*, 1942, 14, 12.

[4] Calculated from Plessner and Richards, *Trans. Faraday Soc.*, 1946, 42 A, 206.

In the case of dipole rotation, the same principles seem to be applicable. The data of Table I show, broadly, that ΔS is more positive, the larger is the molecular (or polymer) unit which must be displaced in the relaxation process. Thus, polyvinyl chloride, the phenol-formaldehyde polymer and the ethylene glycol phthalate polymer, in all of which the molecular units are large, show very high positive values of ΔS, together with high values of ΔE. It is reasonable to assume that for these large molecules, relaxation involves extensive local disorganisation of the lattice, which, as shown above, would account for the energy and entropy changes. In addition, the similar values obtained for polyvinyl chloride + 30 % tricresyl phosphate for the two cases, suggests that the polymer units involved in the two processes have roughly equal dimensions.

The values for dioctyl phthalate in polythene and polystyrene (Table I) are especially illuminating. The former case gives a large ΔE and a considerable, positive value of ΔS; the latter has a small ΔE and a negative ΔS. It is well-known that as compared with polystyrene, polythene is much more close-packed in structure (largely owing to its crystallinity). It can be inferred from this that a given polar molecule dissolved in the solid and rotating in an electric field, will cause much more lattice disorganisation in polythene than in polystyrene; hence the former solid shows a positive ΔS, and the latter a negative one, while ΔE is much greater in the former case. In the field of water vapour diffusion, the same structural difference shows itself in the much larger diffusion constant of polystyrene as compared with polythene.

The occurrence of a *negative* entropy change in polystyrene can be accounted for in a manner analogous to that given for diffusion. It is assumed that only those dipoles are displaced which possess a particular vibrational configuration. As before, since this condition could only become the rate-determining factor when no appreciable energy was required for lattice disorganisation, the negative entropy term would be accompanied by a small ΔE term. Table I shows that this holds for every case, and that, furthermore, there is a linear relation between ΔE and ΔS, so that the larger ΔE, the more positive ΔS becomes. Similarly, the data show that there is a close dependence of ΔE and ΔS upon molecular structure.

The negative entropy values found by Whiffen and Thompson for liquid systems are entirely in accord with these views. In liquids the hole-formation energy is naturally small, owing to the system lacking the rigidity of a solid. This factor, together with the weak intermolecular attractions in the liquids concerned (which are all non-associated) account for the small ΔE values. It seems probable that the unimportance of hole-formation requirements for dipole rotation in these cases, enables the vibrational direction of the dipole to play a determining role in the overall process; this would account for the negative entropy change. Indeed, in non-associated liquids, which on account of the weak intermolecular forces will lack any marked "structure", negative values of ΔS will probably be the rule rather than the exception. In associated liquids, however, positive entropy terms are likely; as evidence for which may be instanced the viscous flow of water, having ΔS equal to about + 5 cal./deg./mole.

The above observations show that it is possible to interpret thermodynamic data relating to relaxation times and diffusion rates in molecular terms. The consistency of the interpretation also confirms the applicability of Eyring's equations to both of these phenomena, more particularly for purposes of comparison. The relationship established between the molecular processes involved in dielectric relaxation and diffusion should also provide a useful basis for clarifying knowledge of the molecular behaviour of solids in respect of other properties.

Summary.

The activation energy and entropy terms associated with dielectric relaxation and water vapour diffusion in solids, are shown to be capable of interpretation in terms of the molecular processes involved. Positive entropy changes are attributed to the occurrence of lattice disordering (hole-formation); negative changes are due to the requirement of a particular vibrational orientation by the molecule involved in either process. The suggested interpretation also accounts both for the observed magnitudes of the activation energy, and for the linear relation between ΔE and ΔS.

Fulmer Research Inst.,
Stoke Poges,
Bucks.

GENERAL DISCUSSION

Dr. Penrose (*Oxford*) said: In view of the interest in the behaviour of dilute solutions at centimetre wave-lengths, Mr. Loubser of the Clarendon Laboratory has measured the dielectric constants and power factors of several non-polar solvents. The measurements were made using a completely filled H_{01} cylindrical resonator. The values obtained at a wave-length of 1·35 cm. are :—

Liquid.	Purity.	ε_{20}(Vacuum = 1).	tan δ.
Benzene . . .	Analar (B.D.H. Ltd.)	2·283 ± ·001	0·0011
Benzene . . .	Analar dried over sodium for one day	2·277 ± ·001	0·0009
n-hexane . . .	Commercial	1·901 ± ·001	0·0008
Cyclo-hexane . .	A$^+$	2·025 ± ·001	0·0001$_5$
n-heptane. . .	A$^+$	1·922 ± ·001	0·0008
Carbon tetrachloride .	Commercial	2·239 ± ·001	0·0008
Carbon bisulphide .	A$^+$	2·648 ± ·001	0·0007

* A$^+$ Purified sample kindly supplied by Dr. H. W. Thompson of the Physical Chemistry Laboratory, Oxford.

Similar measurements at 3·2 cm. wave-length yield power factor values about half as large as these, while the dielectric constants differ by less than 3 parts in 10,000 from those given in the table.

Mr. Whiffen (*Oxford*) said: If a single relaxation time and a Debye loss curve are assumed for a solution of 4g. nitrobenzene/100 cc. then the maximum possible value of the attenuation coefficient, α, under the conditions of the experiments described by Hall, Halliday, Johnson and Walker, is 0·23 assuming the dielectric constant of the solvent to be 2·4 and the dipole moment of nitrobenzene to be 3·95 D. Their Fig. 3 shows that the maximum (as the time of relaxation is varied with solvent viscosity) is for each solvent mixture equal to or lower than this value. The lower values are found in the cases where this maximum occurs at high viscosities and it is likely that this is the result of a broadening of the whole loss curve in solutions of high viscosity. These results agree, therefore, with those of Jackson and Powles and of Whiffen and Thompson ; as a rough rule it may be stated that it is unsafe to assume that the Debye loss curve is obeyed in solutions with a greater viscosity than 1·5 centipoises.

Mr. C. H. Collie (*Oxford*) said: Some measurements have now been made on the attenuation of 10 cm. waves in heavy water. The results obtained are expressed as the ratio of the attenuation coefficient in heavy water to that in ordinary water at the same temperature.

Temperature ° C.	5	20	40	60
$\frac{KD_2O}{KH_2O}$ . .	1·31	1·29	1·26	1·23

The observed ratio is very nearly the same as the ratio of the viscosities and is certainly less than the factor $\sqrt{2}$ which might have been expected on a free rotation theory. The observed ratio is in agreement with the point of view put forward by Professor Bauer in his paper.

Dr. J. G. Powles and Prof. Willis Jackson (*Manchester*) (*communicated*): The table below gives a comparison of the results obtained by Messrs. Whiffen and Thompson, Messrs. Hall, Halliday, Johnson and Walker and ourselves on a range of benzene solutions at wave-lengths around 3 and 1 cm. The figures quoted in column (6) have been obtained by use of a wave-guide technique similar to that described by Whiffen and Thompson, but with the differences that the absorption cell was considerably longer than theirs (60 cm. as compared with 5-20 cm.) and that standing wave detectors were incorporated both behind and in front of this cell. In deriving these figures from the observed differences in power level on the two sides of the cell account has been taken of the effect of internal reflections. It seems that the effect of these may be marked with cell lengths of the order employed by Whiffen and Thompson. The results given in column (6) are not regarded as reliable, however, since certain difficulties associated with the measurements have not yet been satisfactorily resolved. A fuller statement will be made at a later date when the discrepancy between these results and the resonator measurements given in column (5) has been accounted for.

It will be seen that the results obtained by ourselves are, with one exception, consistently higher than those of Whiffen and Thompson at both wave-lengths of measurement, the differences in each case being much greater than can be accounted for by the slight differences in wave-length. There are consequential discrepancies between the dipole moment values derived from the two sets of measurements, though these are masked to some extent by the fact that the dipole moment appears as μ^2 in the relation connecting tan δ and the molecular constants. With the exception of the benzophenone solution, the derived values of relaxation time are, however, closely comparable. The single tan δ value given in column (3) has been calculated from the attenuation data afforded in Fig. 3 of the paper by Hall, Halliday, Johnson and Walker and it would be of interest to have figures from them for other polar molecules in benzene solution.

The results of Table II, Paper I, by Whiffen and Thompson and those given by Hall, Halliday, etc. accentuate the fact brought out by Fig. 4 of our paper that measurements of absorption in solutions employing complex solvents at constant temperature do not afford reliable data regarding molecular constants unless measurements are carried out over a considerable frequency range. Even in the case of simple solvents such as benzene the discrepancies revealed in the comparative table given above emphasise the importance of making measurements at more than two frequencies.

TAN δ VALUES FOR SOLUTIONS IN BENZENE OF CONCENTRATION 1 G. MOLE/100 CC.

1. Solute.	2. Resonator Measurements W. J. & P. at 3·06 cm.	3. Resonator Measurements W. & T. at 3·26 cm.	4. Waveguide Measurements H. H. J. & W. at 3·27 cm.	5. Resonator Measurements W. J. & P. at 1·23 cm.	6. Waveguide Measurements W. J. & P. at 1·22 cm.	7. Waveguide Measurements W. & T. at 1·27 cm.
Benzophenone .	2·64	2·19	—	1·76	1·48	1·29
Nitrobenzene .	4·38	3·58	4·21	3·66	3·08	3·32
Brombenzene .	0·72	0·54	—	0·66	0·64	0·58
Acetone .	0·85	0·67	—	1·69	1·57	1·43

Mr. J. G. Powles (*Manchester*) (*communicated*): It is noticeable that the loss factor for benzene at a wave-length of approximately 1¼ cm., reported by various workers (Jackson and Powles, Whiffen and Thompson, Loubser) is of the same value, namely 0·0011; I have obtained this value for benzene of a variety of standards of purity. It has been suggested that the loss may be due to the presence of water in the benzene and it is interesting to calculate the maximum loss factor to be expected using the known dipole moment of water, assuming the benzene to be saturated and that the Debye relation holds. This gives a maximum loss angle of 0·0026 which is of the right order of magnitude. Using the value of 0·0011 at 1·23 cm., it may be calculated that the maximum of loss factor should occur at 0·27 cm. wavelength, which is again of the correct order when compared with the values for acetone and benzophenone in benzene. Moreover, the calculated values at longer wave-lengths agree reasonably well with the values I have measured at other wave-lengths (except at 23 cm.).

Solute.	Relaxation Time in sec. × 10^{12}.
Water .	1·44
Acetone .	3·3
Benzophenone	16·4

$\tan\delta \times 10^3$	Wavelength in cm.	1·23	3·05	9·1	23·0
	Calculated value with $\tau = 1{\cdot}44 \times 10^{-12}$	1·10	0·47	0·15	0·06
	Observed value	1·10	0·50	0·24	0·14

It would appear, therefore, that the observed loss in benzene is probably due to the presence of water rather than an absorption band in the benzene itself but further measurements under more stringent conditions and at shorter wavelengths are necessary to make a definite decision possible.

Mr. H. Warburton Hall and **Mr. I. G. Halliday** (*Farnborough*) (*communicated*): We agree with the observation of Mr. Whiffen concerning the possible broadening of loss curves for high viscosity solutions, resulting in greater divergence from the Debye curve, since the results of our earlier attempts to use partially polymerised styrene as a high viscosity solvent indirectly support this possibility. Again recent preliminary results of loss measurement, at room temperature, at a lower frequency ($\lambda \simeq 10$ cm.) on the solutions described in our paper suggest a similar broadening effect; the loss maxima, as expected, tending to appear at higher viscosities than for the 3 cm. wave-length measurement. We hope to examine this effect experimentally in greater detail at an early date.

With regard to the remarks of Mr. J. G. Powles and Professor Willis Jackson and their table comparing tan δ values, we give below results of our measurements on the other polar molecules they mention. The substances were of "Analar" quality. The measurement at 3 cm. wave-length were made using the impedance circle method described in our paper which, we believe, contains no errors due to ignorance of internal reflections, etc., and which gives good accuracy. The values of α (attenuation constant) were obtained as described in the paper and β (phase constant) by using a short (0·888 cm. in length) short-circuited cell in the line. From these two constants the values of ϵ'', ϵ' and tan δ given below were derived.

For the 1 cm. wave-length measurements, however, owing to the wave-guide dimensions, it is not easy to apply the impedance method and we have used, therefore, for the time being, an "optical" wave-guide method similar to that used by Messrs. Whiffen and Thompson. We agree that this method is not completely satisfactory. However, an alternative method, at present under investigation, promises to be more satisfactory from the point of view of accuracy than the "optical" method and less laborious with regard to computation than the impedance circle method. In this method, the height of liquid in a vertical guide or, for lower frequencies, coaxial cable is gradually increased and a trace of probe detector output against height of liquid is obtained. The trace is periodic in form with decreasing amplitude and from it α and β are computed.

DIELECTRIC MEASUREMENTS OF BENZENE SOLUTIONS.

Concentration . .	4g./100 cc.		1 g.mol./100 cc.		4g./100 cc.	
Wavelength . .	3.27 cm.			1.23 cm.	•	
Solute.	ϵ'.	ϵ''.	tan δ.	tan δ.	$\mu.^2$	ϵ' at I Mc.
Benzophenone .	2·44	0·127	2·27	1·45	2·200	2·57
Brombenzene .	2·43	0·0375	0·606	0·59	2·253	2·50
Acetone . .	3·00	0·175	0·848	1·58	2·233	3·27

For interest, we have included in the table measurements at room temperature of ϵ' at 1 megacycle measured with a "Dielcometer" and the squares of refractive indices (μ) obtained with an Abbé refractometer.

Dr. J. Lamb (*Manchester*) (*communicated*): Contemporary with the work recorded by Dr. Penrose, measurements of the permittivity and loss-factor values of a variety of solid and liquid dielectrics at frequencies of 10^{10} c/s. and $2{\cdot}4 \times 10^{10}$ c/s. were carried out at Manchester University by utilising a cylindrical resonant cavity operating in the H_{01} mode.

Our analysis relating to this system[1] is rather more general than that described by Mr. Penrose in that we have considered the introduction into the resonator of a test specimen of any arbitrary thickness. In this way it is possible to select the size of test dielectric most convenient for measurement purposes. The consequent flexibility of this method renders it more preferable in cases where the materials under investigation possess widely different loss-factor values. A further advantage arises from the fact that the permittivity and tan δ-value of a particular material can be derived from measurements on the same specimen.

At both frequencies of measurement I have found that the oscillator frequency and amplitude have been sufficiently stable to permit a full resonance curve to be obtained. The Q-factor values, as determined from the measured widths of a derived resonance curve at heights ranging from 0·2 to 0·8 of the maximum, have been found to agree to within $\pm 1\frac{1}{2}$ %. This affords a useful check on the square law rectification property of the particular crystal employed in the detection system.

---

[1] Horner, Taylor, Dunsmuir, Lamb and Jackson, *J.I.E.E.*, 1946, 93, (iii), 53.

Dr. R. P. Penrose (*Oxford*) (*communicated*): In the experiments I described, the dielectric constant was first calculated from measurements on discs of arbitrary thickness. The results of a single experiment are consistent with a series of values of the dielectric constant; thus to determine this quantity *a priori* at a particular frequency one must use at least two different thicknesses of dielectric.

The estimated accuracy quoted in the table of results was generally limited by errors in machining the disc. That the importance of machining errors will vary markedly with the electrical length of the dielectric can be inferred from the accompanying table, which refers to a frequency of 25,000 Mc/s. and a material of dielectric constant 4·0. The system is insensitive to small changes in the thickness of a half wave-length or whole wave-length disc because the air-dielectric interface is at a node of the electric field. Evidently for higher dielectric constants and higher frequencies it will be increasingly desirable to use a specimen of this type for reducing the errors in the dielectric constant measurement.

Thickness of Disc.	Error in Thickness which Corresponds to a 1 % Error in Dielectric Constant (mm.).
λ (= 6·8 mm.)	0·5
$\frac{3}{4}\lambda$	0·03
$\frac{5}{8}\lambda$	0·05
$\frac{1}{2}\lambda$	0·4
$\frac{3}{8}\lambda$	0·04
$\frac{1}{4}\lambda$	0·011
$\frac{1}{8}\lambda$	0·004

The use of a disc whose length is not electrically simple increases the labour involved in calculating the power-factor value from the experimental results. At wave-lengths in the 1 cm. region there hardly seem to be sufficient practical advantages to compensate for this.

Dr. F. C. Frank (*London*) said: I wish to raise objection to a procedure not confined to Professor Jackson, Mr. Collie, or other contributors to this discussion, but common to the majority of physicists and engineers. It consists in writing expressions like

$$E = E_0 e^{j\omega t}$$

when nothing of the sort is meant: the field in a dielectric, or the volts across the terminals of a condenser, are never imaginary.

$$E_0 e^{j\omega t} = E_0(\cos \omega t + j \sin \omega t)$$

represents a rotating vector, the projection of which on a fixed axis represents the oscillating function intended. A correct statement is

$$E = \mathbf{R}(E_0 e^{j\omega t})$$

where **R** signifies " the real part of ". However, on combining this expression with, say, a complex dielectric constant $\epsilon^* = (\epsilon' - j\epsilon'')$ to calculate the electric displacement after the model $D = \epsilon E$, we can no longer use ϵ^* as a simple multiplier, for the result we want is not

$$\epsilon^* \mathbf{R}(E_0 e^{j\omega t}) \text{ but } \mathbf{R}(\epsilon^* E_0 e^{j\omega t})$$

The only consistent interpretation is that we are using an operational calculus, in which ϵ^* is an operator with the operative rule that it acts as a multiplier within the " reality " sign. Alternatively stated, it operates on the rotating vector, changing its magnitude and phase, and the projection of this new vector is the resultant oscillating displacement we require. The fact that by omitting the reality sign, and never seeing this difficulty we make two compensating errors, is no justification.

The point is not merely pedantic or pedagogic. By omitting the reality sign, we not only tend to mislead the student into thinking that he can always throw an imaginary make-weight into a calculation at the beginning, and throw it away again at the end, but we hinder ourselves from combining such calculations with other powerful mathematical techniques in which both the real and imaginary components of the complex exponential are given physical significance. The engineers may, if they wish, retain a private mathematical system for use in limited problems ; but the physicists must be asked to refrain from teaching such methods to the chemists, whose interests are too wide to tolerate a restrictive notation.

Dr. Magat (*Paris*) said : Some time ago, Professor C. P. Smyth, Dr. Schneider and I determined the dielectric properties of bromine and chlorine derivatives of normal aliphatic hydrocarbons at the 10 cm. wave-length with the second Drude method. Since only one point of the dispersion curve was measured, we calculated the critical frequency with the Debye formula according to the shape of the Cole and Cole plots. In view of this uncertainty and of the objections that can be made against the use of the Drude method, we have refrained from publication. But although our absolute value of the critical frequencies may be somewhat uncertain, the general trend is too pronounced to be doubted. This general trend confirms completely the conclusions of Whiffen and Thompson as to the effect of the carbon chain length as soon as the chain length exceeds 8 atoms, in as far as the dependence of the critical frequency on the chain length becomes very small. I may add perhaps, that on the contrary, with shorter chains the critical frequency depends more on the chain length than on the dipole group. The relaxation frequencies of nitro- and chloro-compounds are not very much different (15-25 %) from those of bromine compounds of the same chain length.

Analogous results were observed by Fisher[2] on solutions of aliphatic chlorides, although this author did not draw from his data the above conclusion.

Dr. Thompson (*Oxford*) said : I should like to make a few general comments on this work on the absorption of microwaves by dilute solutions. Many of the papers submitted for this discussion have dealt with theoretical calculations on dielectric loss, and comparatively few experimental results have been reported. While much of the theoretical work is of great interest, and marked progress has been made in interpreting the phenomena with solid dielectrics, it cannot be said that any satisfactory theory has been found to deal with polar liquids and some polar solutions. In such cases the original theory of Debye, in spite of obvious inadequacies, remains perhaps the most satisfactory starting-point.

It seems to me that before much real progress with the theory can be made, more experimental results must be obtained on carefully chosen systems. Measurements of dielectric loss can be of three kinds, namely (1) absolute determinations of the loss of materials which are to be used in practical applications, (2) measurements with polar liquids or solids to test current theories or to discover, empirically, new phenomena such as the occurrence with a given substance of several loss maxima at different frequencies, (3) measurements on the absorption of substances in related series planned in order to build up information on which new theories can be based. Microwave technique has been developed for the most part by radio engineers, whose subsequent measurements have mainly been in group (1), whereas in physical chemistry those of group (3) seem to have received little attention. An apt analogy can be drawn here with another line of work, namely that dealing with the physical properties of electrolytic solutions. The original theory of Debye and Hückel for strong electrolytes involved the basic conception of the ionic atmosphere, and an experimental study of the variation of electrical conductivity, ionic transport numbers and other similar quantities with the viscosity or dielectric constant of the solvent, or valency type of the ions, provided a sound basis for the understanding of how the ionic atmosphere was built up. In the case of dielectric loss with polar liquids or dilute solutions deviations from the original theory of Debye must in general be due to the disturbing influences of the other molecules which surround a given dipole, in other words to the dipole atmosphere. The real problem from an experimental standpoint should be the study of how the properties of a dipole, and in particular its time of relaxation, vary with its particular environment and with other factors such as temperature or viscosity. If more were known about this, we should be in a better position to understand how molecules pack in the liquid state, and which factors operate in dielectric loss. There are indications in the results already available that the time of relaxation of a given polar solute in different non-polar solvents varies with the precise shape of the solvent molecules. Another point which may reveal some of the salient underlying principles is the determination in more cases of the apparent energies of activation for the relaxation process, although the whole picture of relaxation as a rate process requires further examination.

---

[2] *Physikal Z.*, 1939, 49, 645.

Doubts have sometimes been expressed about the value of measurements with microwaves on the loss of dilute solutions, where the theory of Debye may apply satisfactorily. It seems to me that the alternative approach using polar liquids themselves involves the study of many complicating factors simultaneously. With the simpler cases of dilute solutions on the other hand it is usually possible to study the variation of one factor at a time, and the deviations which arise as the solution is made more concentrated can thus be analysed more systematically.

It seems to me therefore that if the excellent technical achievements of the physicists and radio engineers are to be fully utilised in chemistry, the approach should be made rather more in the way which I have just outlined. Only in this way can the complex intermolecular influences be disentangled, and physico-chemical theory developed.

Dr. G. Gee (*Welwyn Garden City*) said: I should like to comment on the negative entropies of activation reported by Whiffen and Thompson in Table V of their second paper. The generally accepted view of the activation process in dielectric relaxation or viscous flow assumes it to involve a loosening of the structure in the neighbourhood of the molecule which is about to respond to the field. If this is so, the energy of activation is employed essentially in producing a local expansion, equal to the volume change $\Delta V^{\ddagger}$ in the activation process. It follows that $\Delta V^{\ddagger}$ can be calculated from $\Delta H^{\ddagger}$ either from the internal pressure $\left(\frac{\partial E}{\partial V}\right)_T$ or from the cohesive energy density $(L_0 - RT)/V_0$ where L_0 and V_0 are respectively the molar latent heat of evaporation and molecular volume. These two quantities are in general nearly equal, and for most liquids lie between 50 and 150 cal./cc. Taking the values of $\Delta H^{\ddagger}$ given in Table V we conclude that, if this energy arises solely from local expansion, $\Delta V^{\ddagger} \approx 20$ cc./mole. A similar argument applies to viscous flow, and in this case it can be independently checked from the effect of pressure on viscosity. It has been pointed out [3] that if the pressure and temperature are simultaneously altered so as to keep the volume constant, the viscosity of a non-associated liquid is also nearly constant. If we write

$$\ln \eta = A + \Delta G^{\ddagger}/RT,$$

we have

$$\left(\frac{\partial \ln \eta}{\partial T}\right)_P = -\,\Delta H^{\ddagger}/RT^2,$$

$$\left(\frac{\partial \ln \eta}{\partial P}\right)_T = \Delta V^{\ddagger}/RT,$$

$$\left(\frac{\partial \ln \eta}{\partial T}\right)_V = 0 = \left(\frac{\partial \ln \eta}{\partial T}\right)_P + \left(\frac{\partial \ln \eta}{\partial P}\right)_T \left(\frac{\partial P}{\partial T}\right)_V.$$

---

[3] Glasstone, Laidler and Eyring, *The Theory of Rate Processes*, p. 486.

Combining these equations, and remembering that, at ordinary pressures,[4]

$$\left(\frac{\partial P}{\partial T}\right)_V = \frac{1}{T}\left(\frac{\partial E}{\partial V}\right)_T,$$

it follows that

$$\frac{\Delta H^{\ddagger}}{\Delta V^{\ddagger}} = \left(\frac{\partial E}{\partial V}\right)_T.$$

In view of the close parallelism between dielectric relaxation and viscous flow, this result must be taken to afford strong evidence for the above estimate of $\Delta V^{\ddagger} \approx 20$ cc./mole. Now if such an expansion were produced by applying a hydrostatic tension it would result in an increase of entropy nearly equal to $\Delta H^{\ddagger}/T$. The entropy of activation will depend to some extent on the distribution of the extra free volume, but it seems to me most unlikely that such a large expansion could be accompanied by a considerable *decrease* of entropy. It must be borne in mind that the entropy of activation is not deduced by thermodynamic reasoning, but is dependent on the accuracy, in an absolute sense, of the statistical rate equation :

$$\frac{1}{\tau} = \frac{kT}{h} e^{\frac{\Delta S^{\ddagger}}{R}} e^{-\frac{\Delta H^{\ddagger}}{RT}}.$$

It seems much more likely that the factor kT/h is in error than that $\Delta S^{\ddagger}$ is actually negative.

Professor E. Bauer (*Paris*) said : It is not surprising to find negative entropies of activation : it only means that the activated complex is in a state of better order than the initial system, *i.e.* it takes less volume in the configuration space.

In the theory which I have given, the relaxation time is proportional to the inverse of the effective length of the potential barrier. If the potential col through which the dipoles can go is very narrow, it is obvious that the entropy of activation must be negative.

Dr. R. J. W. Le Fevre (*Farnborough*): I wish to make three comments on the papers of Thompson and Whiffen.

(1) The figures shown under " μ (calc.) " in Table I (Paper No. 1) are of special interest in those cases where reputable dipole moment determinations *in the gaseous state* have been recorded. It happens that these are all instances for which, when the usual Debye-Clausius-Mosotti equations are used, the apparent moments measured in non-polar solvents are less than the true values. Accordingly—since essentially the same theory underlies the present work—it is at least a satisfactory check on experimental accuracy to note that the moments now deduced from the losses in C_6H_6 solution are all nearer to those found by conventional medium wave dielectric constant methods when this solvent is used, than to those calculated from similar procedures with vapours. Presumably, therefore, even at V.H.F., effects due to association of solvent around solute molecules, are not to be disregarded any more than in the megacycle range. Willis Jackson and Powles (previous paper) report that the tan δ log f curves for benzophenone indicate moments in benzene and paraffin solutions agreeing within 1½ %. Are the graphs capable of demonstrating that μ_{paraffin} is greater than μ_{benzene}?

Substance.	$\mu_{C_6H_6}$.	$\mu_{T. \& W.}$	μ_{gas}.
Ethyl benzoate .	1·92	1·85	1·95
Nitrobenzene .	4·05	3·75	4·23
Chlorobenzene .	1·54	1·52	1·73
Bromobenzene .	1·55	1·51	1·71
Toluene . .	0·34	0·30	0·37
Nitromethane .	3·02	2·87	3·54
Acetone . .	2·76	2·58	2·84

(2) Whiffen and Thompson's discussion of relaxation as a rate process is most suggestive. *A priori* it seems wrong to attempt to connect the macroscopic viscosity with an essentially molecular movement. Following

---

[4] Guggenheim, *Modern Thermodynamics*, equation (155), p. 35.

Frenkel's treatment of liquid benzene, the temperature coefficients of viscosity at the temperatures of observation would appear more likely to yield a correlation, since these should reflect the ability of applied forces (and thus parallel applied alternating electric fields) to overcome intermolecular orienting influences. Unfortunately such an approach will require a great deal of new experimental work.

Nevertheless it may be worth pointing out that the sequence based on τ values in solution of compounds in Table I holds roughly for a number of simple properties of the pure substances, where these properties are obviously some measure of intermolecular attractions. As instances, the coefficients of cubical expansion and b.p.'s may be quoted :—

Substance.	Ph_2CO.	$PhCO_2Et$.	$PhCO_2Me$.	$PhNO_2$.	PhBr.	PhCl.	PhMe.	$MeNO_2$.	Me_2CO.	$MeCO_2Me$.
10^4 × (cc./degree) .	—	88	88	83	156	98	110·9	119	143	—
B.p. (° C.) . .	306	213	199·5	211	92	132	111	110	56	57

As $T_{\text{B.P.}} \approx 2/3\, T_C$, the critical temperatures might provide a smoother series. Not many have been ascertained experimentally, the following, however, are relevant :

PhBr, 670°; PhCl, 632°; PhMe, 594°; Me_2CO, 508°; $MeCO_2Me$, 505·9°.

Thus acetone and methyl acetate are put into the observed order by critical temperatures but not by b.p.'s.

(3) Thompson and Whiffen, in Table II (Paper No. 2), calculate a moment of 0·19D. for *p*-cymene, and comment that the literature value (0·29) is unreliable. Le Fèvre, Le Fèvre and Robertson [5] have also criticised the earlier dielectric constant data for this hydrocarbon. A full investigation of its polarisation both in solution and as a pure liquid, showed this to be 45·6 to 45·7 cc. $(R_L)_D$. being 45·2 cc., the moment was therefore *ca.* 0·15D.

Normally such small μ's would be very undependable but with *p*-cymene an independent check is available from its electric double refraction compared with that for benzene and toluene for which, respectively, $R = 15{\cdot}7 : 12{\cdot}1 : 24{\cdot}3$. Since, for molecules whose resultant moments lie along the axis of maximum polarisability, the Kerr constant is the sum of an anistropy term and a dipole term, and since the anistropy order is obviously *p*-cymene > toluene > benzene, the only reasonable conclusion is that the moment of cymene is less than that of toluene, *i.e.* between 0·37 and zero. Thompson and Whiffen's figure (0·19) thus seems quite satisfactory.

Mr. J. G. Powles and Professor Willis Jackson (*Manchester*) (*communicated*) : In reply to Dr. R. J. W. Le Fevre our results seem to indicate that if there is any difference between μ paraffin and μ benzene it is less than 3 % ; since some of the measurements are only accurate to ± 10 % a difference in the areas under the two curves of about 6 % would be required to definitely establish a difference.

Dr. F. C. Frank (*London*) (*communicated*): Dr. Le Fèvre's agreement provides very satisfactory evidence of the accuracy of measurement, but it is no indicator of the correctness or otherwise of theories of solvent effect or " internal field ". It is a consequence of any linear theory of dielectrics that the same result will be found, when the estimate is made, as by Jackson and Powles, from the integrated area under the loss curve. When the estimate is made by Whiffen and Thompson's method, there is also a check on the correctness of the simple Debye form of the loss curve, but still no possibility of discriminating between, say, Debye's and Onsager's theories of internal field. Different pairs of values will be calculated according to the two theories, but still in agreement provided the measurements are correctly made.

---

[5] *J. Chem. Soc.*, 1935, 180.

Dr. H. Pelzer (*Greenford*) said: Whiffen and Thompson in their second paper have raised the question whether their formal result of a negative entropy change by using Eyring's way of writing

$$\frac{1}{\tau} = \frac{kT}{h} e^{\Delta S/R} e^{-\Delta H/RT}$$

has an absolute physical significance. Now the expression for an absolute rate process can also be written in a more explicit form containing a number of terms each of which corresponds to one degree of freedom of the initial or the transition state. Quantum effects need be considered only with reacting bodies as light as hydrogen, whilst with polar groups classical treatment is sufficient. In the classical case the factors mentioned above can be written down [6] without involving Planck's constant *h*. Eyring apparently prefers to normalise the partition functions by expressing them in terms of *h* so that his expressions apply equally to classical and quantum [7] treatments. Apart from this, Eyring groups all degrees of freedom of the transition state except one together with all degrees of freedom of the initial state whilst the remaining degree of freedom of the transition itself is split off and incorporated into his first factor kT/h. This provides an elegant way of writing down theoretical rate equations and is convenient for comparing different rate processes in a standard manner; and it so happens that the order of magnitude of this factor is of the order of the collision frequency under normal conditions. But it leaves in the remainder of the expression an unbalance of one dimension and therefore the sign of Eyring's ΔS has no absolute significance; for by chosing the units of phase space greater than *h* (which is permissible in classical cases) one can always convert a negative ΔS into a positive one.

It is a different matter if the ΔS of different processes are compared and one found to be negative or more negative than others. This case means plainly that a reaction rate turns out to be slower than expected. In this connection mention must be made of a paper by Kramers [8] who has made a critical investigation of the transition state method in terms of diffusion and Brownian motion and the frequency of collisions with the temperature bath, described formally by a quantity η. Though for a large range of η-values the transition state method gives correct results, Kramers finds that for very small and for very large values of η it predicts too high reaction rates. In the case of small η the assumed Boltzmann equilibrium between initial and transition state can not be maintained, whilst for very large η there is an increasing probability (neglected by the transition state method) of the system diffusing back after it has just passed the barrier. Kramers uses Smoluchowski's equation (*i.e.* a diffusion equation augmented by a term taking into account an external field of force) which may be of importance not only for the total rate of reaction as treated by Kramers but also for the problem of time constant distribution by providing a more rigorous treatment than the simplified assumptions of the transition state method. It must be borne in mind, however, that even a small variation of barrier height about an average value will have a much bigger effect on the distribution of time constants than the introduction of the more rigorous method just mentioned.

Mr. Whiffen (*Oxford*) said in reply to Dr. Gee: I agree that the question of the applicability of the absolute reaction rate theory equation to relaxation rate phenomena is an open one. If, however, negative values of the entropy changes of activation can be taken at their face value a simple qualitative picture can be proposed. In this, molecules neighbouring on that which is rotating must pack themselves closely together in order to allow the central molecule to move ; the total decrease of entropy of these neighbours might be greater than the increase of the rotating molecule itself, so that the process as a whole results in a diminution of entropy. The larger the central molecule, the greater is the degree of packing required and therefore the more negative the entropy change ; this reasoning agrees with the data for solutes dissolved in *n*-heptane given by Whiffen and Thompson in their second paper, Table V.

---

[6] *Z. physik. Chem.*, *B*, **15**, 467, 1932, eq. (23).

[7] Wigner, *Z. physik. Chem.*, *B*, 1932, **19**, 203.

[8] *Physica*, 1940, **7**, 284.

In this connection it may be pointed out that variable temperature measurements at constant volume, as opposed to constant pressure, would be expected for simple qualitative reasons to lead to smaller activation energies and thence to even more negative values of the corresponding entropies of activation. In conclusion I would like to thank Dr. Everett for discussions on this subject.

Mr. Sutton (*Oxford*) said : It is difficult to keep dry benzene dry. This substance is very hygroscopic, having a strong tendency to pick up a small but appreciable quantity of water. The observed residual loss might be due to this and could only be prevented by drying the benzene very carefully and then taking stringent precautions to avoid the ingress of moisture.

Dr. J. H. van Santen (*Eindhoven*) said : The interaction energy of benzene with polar substances is relatively large compared with other hydrocarbons. This strong interaction may be explained in two different ways.

In the first place we may assume that the polarisability of the benzene molecule is high as a subsequence of a low excited level. This hypothesis, however, is not confirmed by the low permittivity of benzene, nor is it able to explain the fact, that chloroform, a molecule with an active positive surface (probably the H atom) gives rise to a positive heat of mixing with benzene, whereas ether, a molecule with an active negative surface (the O atom) gives a negative heat of mixing. The suggestion that the polarisability caused by the Π-electrons of benzene increases with the electric field strength does not give an explanation for the asymmetry in the heats of mixing.

In the second place it might be assumed that in the benzene molecule permanent partial dipole moments occur, which give rise to a large interaction energy with suitable partners such as $CHCl_3$. This second assumption is confirmed by investigations of boiling points. It is not limited to aromatic compounds but might be applied to other double bond containing substances.

Dr. J. Weiss (*Newcastle-upon-Tyne*) (*communicated*) : Recently [9] the mean dielectric constants of a number of molecular compounds and of their respective components in the solid state have been measured in a new type of cell suitable also for small quantities of materials. This has furnished strong evidence in favour of an electron transfer mechanism in the formation of certain molecular compounds, which is also supported by some other properties of these compounds.

Dr. C. J. F. Böttcher (*Rotterdam*) said : Dr. Thompson has remarked that it is necessary to do more experimental work on the subject of the dielectric constant to be able to get a satisfactory theory. Although I support this suggestion I think that another thing is of the same importance, *viz.* to check the existing theories with the aid of measurements that have already been done. In my remarks on the papers of Dr. Frank and Dr. Fröhlich I pointed out already that my objections to these papers are that they hardly consider the numerous experimental data already known. Using these data it appears that the simple model of a molecule in a spherical cavity used first by Onsager leads to unexpectedly accurate results ; with its aid, dipole moments can be calculated from data of pure dipole liquids and in the case of mixtures of two dipole compounds it is even possible to calculate both dipole moments with adequate accuracy.[10]

---

[9] Knonberger and Weiss, *J. Chem. Soc.*, 1944, 464.

[10] Böttcher, *Rec. trav. chim.*, 1943, **62**, 119.

There is another way to show that the simple model mentioned above needs hardly any correction. Starting from this model I derived formulæ for dipole free substances and showed that with the aid of these formulæ all deviations from the Clausius-Mosotti and Lorentz-Lorentz relations can be accounted for.[11] It is therefore important to understand theoretically how it is that the simple model of a molecule in a spherical cavity of a homogeneous dielectric is much more accurate than would be expected on first consideration.

Dr. A. Schallamach (*Welwyn Garden City*) (*communicated*): Professor Evans mentioned that the factor kT/h in the fundamental rate equation, on which the calculation of entropies of activation depends,* may be found to be different in the case of rotational motion as encountered in dielectric relaxation. I carried out the calculation some time ago, using as model a rotator with its mass concentrated at the end of a radius at right angles to a fixed axis of rotation. If α (radians) is the length of the reaction co-ordinate in the activated state the partition function of the activated complex is found to be

$$F = (2\pi IkT)^{\frac{1}{2}}\alpha/h$$

where I is the moment of inertia of the rotator; and the average angular velocity in the activated state is

$$\bar{\dot{\phi}} = (kT/2\pi I)^{\frac{1}{2}}$$

It will be seen that only the moment of inertia enters into these two equations. The life-time of the complex is given by

$$\tau = \alpha/\bar{\dot{\phi}}$$

and the rate constant k becomes again

$$k = (kT/h)e^{-\Delta F^{\ddagger}/RT}$$

as in the case of translational motion.

Mr. Whiffen and **Dr. Thompson** (*Oxford*) said: Since our papers were written, measurements have been made at other wave-lengths, namely 1·65 cm., 0·9 cm. and 0·6 cm. The experimental method used at 1·65 and 0·9 cm. was similar to that described for 1·27 cm. We are much indebted to Dr. J. H. C. Griffiths of the Clarendon Laboratory, Oxford, for the oscillators used. At 0·6 cm. the waves were generated by means of a 1·2 cm. Klystron oscillator working with a crystal frequency doubler; in this case, the original waves of 1·2 cm. length are not propagated since the dimensions of the wave-guide used are below the cut-off value. The absorption cell was a U-tube guide, and the total power lost was plotted against length of cell filled. A small scatter of points about a straight line was obtained, and the best fit was determined by the method of least squares. The scatter was due in part to experimental error, but also partly to a variation of the power reflected with the particular length of cell used. Since a number of different cell lengths were used at random up to a length of some twenty wave-lengths, the slope of the line should give an accurate measure of the attenuation coefficient. The intercept corresponds to an extrapolated value of the loss at zero filling, and this will include the loss by the cell walls, by reflexion and by any transformation of the mode at the slightly curved interfaces. (The transformed energy will not reach the detector since modes other than H_{10} are not propagated by the empty guide.) This method of measurement was necessary at 0·6 cm. since a standing wave detector was not available.

The results are shown in the following table, which gives the molecular loss tangents ($\tan \delta/c$) at 20° C. Measurements at five wave-lengths are included for solutions of chloroform, camphor, and α-bromonaphthalene in *n*-heptane, and for liquid toluene. These losses have been compared with a single Debye loss curve by obtaining the best fit for the two parameters μ and τ. The best fit is obtained with the values of μ and τ shown in columns 2 and 3, and the extent of agreement is shown by comparing the values of the loss to be expected from the Debye curve with those found experimentally. The agreement is in every case within the experimental error of about 5 %. In the figure, the experimental points are shown against the Debye curve calculated for the assumed values of μ and τ. It is interesting to notice that the time of relaxation of chloroform in normal heptane, namely $3·1 \times 10^{-12}$, is less than half the value found by Jackson and Powles for the same solute in benzene. In cyclohexane a value of $2·5 \times 10^{-12}$ was found in the work we have described, although this may be somewhat inaccurate.

---

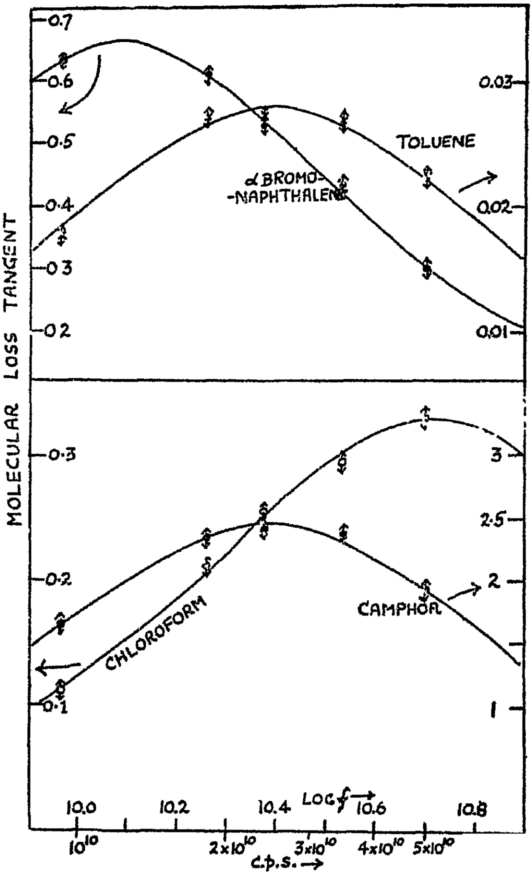
[11] Böttcher, *Physica*, 1942, **9**, 937, 945; *Rec. trav. chim.*, 1943, **62**, 325, 503; 1946, **65**, 14, 19, 39, 50.

* And which is deduced for translational motion.

TABLE I.—Dipole Moments in Debye Units; Values of τ in Units 10^{-12} sec.

Substance.	μ ass.	τ ass.	Values of Loss at Wavelengths (cm.).									
			3·26		1·65		1·27		0·88		0·60	
			calc.	obs.	calc.	obs.	calc.	obs.	calc.	obs.	calc.	obs.
Chloroform	1·09	3·1	0·113	0·112	0·207	0·212	0·251	0·255	0·304	0·296	0·328	0·330
Camphor	2·99	6·5	1·62	1·65	2·36	2·35	2·45	2·42	2·32	2·40	1·95	1·93
α-bromo-naphthalene	1·55	13·0	0·636	0·638	0·610	0·614	0·538	0·535	0·422	0·434	0·300	0·304
Toluene	0·32	6·4	0·0184	0·0175	0·0268	0·0278	0·0279	0·0268	0·0265	0·0269	0·0224	0·0222

In view of the points raised by Jackson with regard to the experimental methods, it is clear, as explained in our papers, that care must be exercised in assessing the possible error due to incorrect allowance for reflexions at the interface of the cell, but such errors could not in our measurements lead to errors in a measured value of tan δ of say 0·02 of more than about 0·001. The main interpretations of our results could not be significantly affected.



With regard to the differences in nomenclature and terminology between the different workers, we welcome the exposition of the inter-relations which Jackson has given.

Mr. D. M. Ritson (*Oxford*) (*communicated*): I should like to comment on Dr. Thompson's remarks. He claims that there is no satisfactory theory of the " time of relaxation " for pure liquids, and that more information on the liquid state will be provided by a study of solutions rather than pure liquids.

The Debye expression for the " relaxation " of a dielectric is really a particular case of a more general law of the same mathematical form. All the many and varied approaches to the problem involving " co-operation," " strong " and " weak " collisions, different dielectric theories, etc., have ultimately reduced themselves to the Debye expression. It would seem fairly certain that any deviations from this law must indicate actual differences in the molecular environment of the dielectric existing for periods of the order of 10^{-10} sec.

The study of pure liquids can be undertaken with the following purposes in view:

(1) Deviations from the Debye theory (*i.e.* " spread of relaxation times ") should provide qualitative evidence on the occurrence of association or aggregation of molecules in the liquid. It may prove possible to decide whether X-ray data are due to the liquid being composed of a very large number of almost perfect micro-crystalline aggregates (as is envisaged by Professor Bauer), or represent a statistical degree of order and nothing more. (2) A very close correlation exists between viscosity, diffusion and the " time of relaxation ". The exact quantitative correlation between these quantities should provide the basis for obtaining an improved theory of liquid structure.

I personally take the rather partisan view that the study of dilute solutions is not likely to provide us with very much more than the information than has been obtained, or could be obtained, from diffusion measurements on solutions (internal viscosity), whereas the study of pure liquids is likely to provide us with a valuable new means of finding the correct theoretical approach to the liquid state.

Dr. Sutherland and **Mr. Cripwell** (*Cambridge*) (*communicated*): It is important in the development of a new field to compare results obtained by various methods and various observers. In the measurement of dielectric losses of dilute solutions of organic materials, Willis Jackson and Powles have used a resonating cavity method, while Whiffen and Thompson have used a wave-guide at 1·25 cm. and a resonating cavity at 3·26 cm. When the results given in their respective papers are reduced to the same theoretical concentration of 1 mole per 100 c.c., the following comparison emerges :—

VALUES OF TAN δ BY DIFFERENT OBSERVERS.

		Benzophenone.		Nitrobenzene.		Acetone.		Bromobenzene.	
Willis, Jackson and Powles	λ in cm.	3·06	1·23	3·06	1·23	3·06	1·23	3·06	1·23
	tan δ	2·64	1·76	4·39	3·66	0·854	1·69	0·724	0·663
		(2·25)	(1·43)	(4·14)	(3·47)	(0·778)	(1·68)		
Whiffen and Thompson	tan δ	2·19	1·29	3·58	3·32	0·67	1·43	0·54	0·58
	λ in cm.	3·26	1·27	3·26	1·27	3·26	1·27	3·26	1·27
Percentage difference		28	17	9	19	15	21	12	25

It will be noted that the differences are considerable, the discrepancy varying from about 10 to 30 %. This seems far outside experimental error. Moreover the differences are all in the same direction and would appear to be too large to be accounted for by the small differences in wave-length used by the different observers. We have recently measured three of the above four liquids using a wave-guide method which we believe to be more accurate than the wave-guide used by Whiffen and Thompson. Our results are included in brackets between those of Willis Jackson and Powles and Whiffen and Thompson. It will be osebrved that our values are intermediate between those of the other two laboratories. It is clear that further investigation is required into the relative accuracy of the various methods now in use. These differences were obscured by the chance agreement between certain values of μ and τ obtained by different observers and this emphasises the need for accurate determination of losses at more than two wave-lengths if reliable values are to be obtained for μ and τ.

Dr. Magat (*Paris*) (*communicated*) : We are perfectly aware that the free energy of activations for viscous flow and for dielectric relaxation are not identical in the mathematical sense of the word. As a matter of fact, it was already pointed out by Eyring and Powell[12] that the first process requires some extra work probably " in order to move the associated liquid or large hydrocarbon from one equilibrium position to the next, work which is not needed for dipole rotation ". The difference is usually of the order of 0·5-1·0 kcal., which, in the actual state of the reaction rate theory, is considered as compatible with the assumption of identical mechanisms. Characteristically, the difference between the two activation energies is the largest for molecules with 8-20 chain atoms,[13] from which we know[14] that they constitute some kind of transition compounds, able to move either by a segment mechanism or as a whole. In the test of the theory advanced by Dr. Schallamach in his discussion remark, the error in V, or in the τ/η ratio, is solely determined by the *absolute* value of the difference of the two free energies of activation and not by its *relative* value, which seems to us more significant. The effect of the difference is further amplified by the fact that this difference enters as an exponential.

The above mentioned difference of 1 kcal. would produce an error in V by a factor of 10. In our calculation, however, the error induced would be of 12-15 % on the chain length of the moving segment, error that seems to us insignificant.

The weakest point in our argument is, in our opinion, not the implicit identification of the free energies of viscous flow and of dielectric relaxation, but the semi-empirical numerical factor 2·45 in our equation (3). Most probably this factor is somewhat too low and the correct one would be somewhere between 2·5 and 5, which would lead to segment lengths between 12 and 30. I think we may discard for the time being the possibility of this factor being widely different for different kinds of rubber.

I may perhaps add, that the results published by Dr. Schallamach since the meeting[2] rather confirm our estimate, the values given by him

Compound.	C_2H_5OH.[1]	$C_6H_{11}OH$.[1]	$C_{10}H_{17}OH$.[2]	$C_{30}H_{49}OH$.[3]	Butaprene NM.	Neoprene.	Hevea.
$E_{\tau}^{\ddagger}$, 25° C.	6·3	7·4	10·1 [15]	13·7 [16]	14·4	14·3	9·5
$F_{\tau}^{\ddagger}$, 25° C.	2·9	4·1	5·5 [10]	6·0 [17]	8·1	7·3	7·3

---

[12] Eyring and Powell, *Advances in Colloidal Science*, New York, 1942, Vol. I, p. 219.

[13] Schallamach, *Trans. Faraday Soc.*, 1946, 42, 495.

[14] Magat (*this Discussion*).

[15] Extrapolated.

[16] At — 27° C., calculated from Schallamach's data.

[17] At — 12° C., calculated from Schallamach's data.

for the activation energy for compounds with 8-16 chain carbon atoms being of the same order as the ones we found for polar elastomers and for smoked sheet, both values being substantially higher than the values found for shorter chain length alcohols, as can be seen from the above table.

Dr. M. Magat (*Paris*) (*communicated*): Small and even slightly negative values for the activation entropy have been already encountered by Eyring and Powell.[1] They were interpreted by them as indicating the possibility of a free rotation even in the inactivated state. This explanation holds probably also for the cases observed by Whiffen and Thompson. However, the absolute values of the entropy they found are larger than the ones mentioned by Eyring and Powell.

E. APPLICATIONS, INCLUDING LIQUID SYSTEMS.

ON THE APPLICATION OF DIELECTRIC MEASUREMENTS TO CHEMISTRY.

By L. E. Sutton.

Received 18*th March*, 1946.

There are two ways of approaching the subject of dielectrics in relation to chemistry. One is to ask what chemistry can contribute to the theory of dielectrics; the other is to ask what a study of dielectrics can contribute to chemistry. In the present discussion, a majority of contributors is likely to stress the first question; so it is well that at least one should stress the second. Only a general survey of the applications will be attempted; and the aim will be to indicate possible developments, not to give a detailed catalogue of what has been done.

Polarisation in a homogeneous material may be considered to arise in three ways (i) by the deformation of electron distributions round atoms—called electron polarisation, (ii) by the displacement of charged atoms from their normal relative positions—called atom polarisation, (iii) by the change in average orientation of permanently dipolar molecules—called dipole orientation polarisation. The chemist is interested in all three of these quantities; so the applications of dielectric measurements may conveniently be considered under these as headings.

Electron Polarisation.—The correct expression for the electron polarisation of a low pressure gas can be written in the Sellmeier form[1] and in this way the concept of an effective movable electron cloud charge ($\bar{e}$) and an effective restoring force constant ($\bar{k}$) can be arrived at, for at low field frequencies (all radio frequencies) the expression reduces to $\epsilon - 1$ or $n^2 - 1 = 4\pi N\bar{e}^2/\bar{k}$. When the gas is not dilute, or for liquids, the mutual effect of the molecules may be taken account of by the cavity field which leads to the Clausius-Mosotti or the Lorentz-Lorenz expression and thus to the definition of molecular electron polarisation or molecular refractivity.[2]

This quantity is usually obtained from refractive index or dielectric constant measurements. In principle it can be calculated from the wave-function for the system,[3] but this has so far been done only for very simple systems. For molecules it shows scalar additivity: the value can usually be expressed very well as the sum of terms appropriate to atoms or, better, to bonds.[4, 5] It is not very sensitive to environmental or constitutive effects; though there are well-known exceptions, *e.g.* the effect of non-localised bonds. Any interaction of bonds should have some such effect; but it is usually small. Because of this fact and because absolute values cannot be easily predicted, this quantity is only of limited usefulness as a means of investigating molecular structure. It is, however, of great value in discussing electron distributions and interatomic forces. Ionic polarisabilities may be used in treating the balance between electrovalent and covalent character in a bond.[6] It is, also, relevant to the tendency of an atom to form dative bonds, to the effect of the presence of a polar group in a molecule on the equilibrium character or the reactivity of another,[7] and to the calculation of intermolecular forces, not only between ions,[8] or between dipolar molecules.[9] but between molecules in general because it enters as a parameter in the treatment of dispersion forces.[10] It is also required in discussing the scattering of radiation by matter; and it was used in treating the energy of non-penetrating orbits.

---

[1] See, *e.g.* Van Vleck, *Theory of Electric and Magnetic Susceptibilities*, Oxford, 1932, pp. 42, 361.

[2] See [1], also Debye, *Polar Molecules*, New York, 1929, chap. 1; and Smyth, *Dielectric Constant and Chemical Structure*, New York, 1931, chap. 1.

[3] See, *e.g.* Pauling and Wilson, *Introduction to Quantum Mechanics*, New York, 1935, p. 227.

While it is, therefore, a chemically useful quantity, it could be made more of one. In particular we need more information about the anisotropy of the polarisability of a molecule, and of bonds too. This is likely to be of considerable importance in the study of detailed mechanism of reaction. The energy and the configuration of activated complexes have been considered largely as arising from the displacement of atoms from their normal places in the reacting species; but in some reactions, especially those involving a polar reactant, the distortion of the electron distribution may be vital, and to treat this it is obviously necessary to have the above-mentioned data. In any case, if the polarisability is anisotropic, the dispersion forces will be also; and these must have some effect upon the energy of approach, especially of large molecules. For understanding the transmission of charge displacements through a molecule, such information is vital. Progress has been made in getting it both for molecules and for bonds.[11, 5] For most single bonds examined the ratio given for the transverse to the longitudinal polarisation is from 0·55 to 0·75; but for the H—H and the N—H bonds it is given as more than one (1·31, 1·45 respectively), and for C—C as 0·01. Double bonds appear to be proportionately more polarisable along their length (ratios *ca.* 0·37); while triple bonds are intermediate. The improbable value ascribed to C—C suggests that further work is needed before these values can be relied on.

One other point, which so far appears not to have been considered quantitatively, is the variation of polarisability with field strength. The electric field in the near neighbourhood of a dipole can be large. At 3 A. axially from a mathematical dipole of moment 3 D. it is 67×10^6 volts/cm.; while at 3 A. axially from the midpoint of a doublet (which is a better approximation to a real dipole) with 2 A. separation between the poles, and moment 3 D., it is $84{\cdot}4 \times 10^6$ volts/cm. The energy of a dipole of moment 1 D. in such fields is 3-4 kcal./mole. It is therefore possible that the polarisabilities obtained by ordinary macroscopic fields

---

[4] Smyth, *op. cit.*, chap. viii.

[5] Denbigh, *Trans. Faraday Soc.*, 1940, **36**, 936.

[6] See, *e.g.* Rice, *Electronic Structure and Chemical Binding*, New York, 1940, chap. xii.

[7] See, *e.g.* Hammett, *Physical Organic Chemistry*, New York, 1940, chap. iii and vii; also Watson, *Modern Theories of Organic Chemistry*, Oxford, chap. vi, also Wheland, *The Theory of Resonance*, New York, 1944, chaps. vii, viii.

[8] See Sherman, *Chem. Rev.*, 1932, **11**, 112; Rice, *op. cit.*, p. 229.

[9] Debye, *Physik. Z.*, 1920, **21**, 178; 1921, **22**, 302.

[10] London, *Trans. Faraday Soc.*, 1937, **33**, 8; *J. Physic. Chem.*, 1942, **46**, 305.

[11] Stuart, *Molekulstruktur*, Berlin, 1934, chap. vi.

will not always be applicable to the discussion of short-range intermolecular forces. There are, indeed, phenomena which indicate this. The polarisation of molecules like ethylene dichloride is abnormally high.[12] The solubility of polar compounds (*e.g.* water, poly-nitrobenzenes or ethylene dicyanide [13]) is greater in benzene than in hexane or carbon tetrachloride despite the similarity of the dielectric constants; and quite stable compounds, even, are formed between poly-nitrobenzenes and aromatic hydrocarbons. Briegleb [14] has attributed this to the shape of the aromatic framework, *i.e.* a flat ring, and to the rather disperse distribution of the electrons in the $\pi\pi$ bonds, which together give approaching polar molecules ready access to polarisable material. But while this is undoubtedly part of the story, it is not certain that it is all; especially since linear unsaturated hydrocarbons show similar behaviour (*e.g.* the addition products of ethylene and certain ions, the high mutual solubility of olefines and liquid sulphur dioxide). A possible inference is that the unsaturated or aromatic nature gives rise to abnormally high polarisability in very strong fields, and indeed this qualitative idea has been current for some time. The whole question of the variation of polarisability with field strength is being examined theoretically in collaboration with Dr. C. A. Coulson. The results will be communicated later.

Atom Polarisation.—As with electron polarisation, atom polarisation can be regarded as arising from the interaction of the alternating field with polar oscillators; and the molecular atom polarisation can be written as the sum of terms $4\pi Ne^2_i/9k_i$, where e_i and k_i are the movable charge and the restoring force constant for the *i*th linear oscillator. But in this case the oscillators correspond much more closely to physical reality, for there are oscillations, relative to one another, of atoms bearing effective charges, *e.g.* in the longitudinal vibrations of dipolar bonds, and in the bending of pairs of dipolar bonds. The fundamental frequencies for these in small molecules occur in the infra-red region; though some of those of large molecules may fall into ultra-high-frequency radio ranges. Usually, then, k is large enough, and e is small enough, to make the atom polarisation small compared with the electron polarisation (5-10 % thereof); but exceptions are known. Atom polarisation can be considerable when large moments, 2·5 D. or more, are held opposed by force constants of ordinary magnitude, as in *p*-dinitrobenzene, mercuric chloride, *p*-benzoquinone and similar compounds; [15] or when the moments are not especially large but the force constant is unusually small, as in the ethylene dihalides.[16] The latter type of polarisation was first discussed as the orientation polarisation of a molecule which acquires a transitory dipole moment when it is bent. But atom polarisation and this bending orientation polarisation are statistically indistinguishable.[16] If the force constant and, hence, the frequency of vibration are large, the main effect of the field is to shift the mean position about which vibration occurs; but as they grow smaller, so its effect becomes more nearly the orienting of a bent, and therefore polar, molecule.

Atom polarisation may therefore be of diagnostic value to the chemist. For example, aluminium trimethyl is partly polymerised in the vapour phase, probably to a dimer. The structure of the polymer, and the reasons for its existence, are uncertain. It is possible that the Al—C bonds are quite highly polar (magnesium dimethyl appears to be salt-like [17]) and that this is the cause of the polymerisation.[18] Should this be so, the

[12] Mizushima, Morino, and Higasi, *S.P.I.P.C.R. Tokyo*, 1934, **25**, 159.

[13] Bloom and Sutton, *J. Chem. Soc.*, 1941, 727.

[14] Briegleb, *Z. physikal. Chem.*, *B*, 1935, **31**, 58.

[15] Coop and Sutton, *J. Chem. Soc.*, 1938, 1269.

[16] See *Ann. Reports*, 1940, **37**, 57.

[17] Krause and von Grosse, *Die Chemie der metall-organischen Verbindungen*, Berlin, 1937, pp. 110, 192.

[18] Skinner and Sutton, *Nature*, 1945, **156**, 601.

monomer and, still more, the polymer would be expected to have large atom polarisations. Wiswall and Smyth[19] found that the total polarisation of the vapour is larger than the electron polarisation expected for the dimer itself, and therefore larger than that for the mixture of monomer and dimer. They supposed that a large atom polarisability is only a remote possibility, assumed a non-polar structure for the dimer, and attributed the excess polarisations to the monomer being pyramidal and polar. This hypothesis does not explain the existing polarisation data perfectly, but nor would the supposition that both monomer and dimer have constant polarisation, *i.e.* large atom polarisation. The question therefore remains open.

When the magnitude of the effective movable charge is known, *e.g.* when the moments of two opposed, movable dipoles are known, the force constant can be calculated; and if, further, the mode of movement may be guessed and the reduced mass thus calculated, its frequency may be derived, as was done for *p*-benzoquinone.[15] The fundamental absorption of ethylene dichloride would thus be expected to be at about 0·2-0·3 mm. Since the vibrations would be notably anharmonic, harmonics should be strong; moreover each "line" would be a complex because there would be transitions from the lower excited levels which would be filled thermally. Finally there would be much broadening by the Stark effect save in gases at low pressures. Direct investigation of such vibrations by absorption or refractive index measurements is much to be preferred; but little work has been done in this difficult region.[20]

Conversely, of course, if the atom polarisation and the force constant are known, the effective movable charge can be calculated. This provides one possible means for determining individual bond moments. The atom polarisation of a crystal might be used to examine the polarity of the bonds in it, especially if the frequencies of vibration of the lattice could be found.

One matter which has not yet been much considered is the relation of atom polarisation and dielectric loss. This is, perhaps, an application of chemistry to dielectrics; but we may turn from our main preoccupation to give it some attention. Clearly, if any of the vibration frequencies were low enough to fall within the region wherein dielectric constants can be measured, they would give rise to a dielectric loss. It is not to be expected, however, that there would be sharp absorption, for there would probably be a multiplicity of levels for the reasons mentioned in connection with ethylene dichloride. So the dielectric constant of a solution or a pure substance would probably fall gradually with increasing frequency, as it does when dipoles cease rotating. We have seen also that the field may produce a rotation as well as an effect on the internal configuration of the molecule. There could, therefore, be dielectric loss by the cessation of this rotation. Even in a low-pressure gas the rotational absorption would be complex, because there would be many moments of inertia, for the many configurations. Ethylene dichloride should show the effect; and so, even, might long chain paraffins, as has been suggested by Mr. R. B. Richards (*private communication*).

Although methylene groups are slightly polar, the paraffins are non-polar as a whole, whatever their configuration, provided that (i) all the C—H bond moments are equal, (ii) all the C—C bond moments are zero (iii) the valency angles H—C—H, C—C—H, and C—C—C are all 109·5°. Mere rotation about C—C bonds as axes will not give rise to polarity. Atom polarisation could arise, however, if the valency angles were changed by an external field. For example, a paraffin initially in its linear staggered configuration (Fig. 1), *i.e.* that of lowest potential energy, could be made polar if the methylene groups were lined up with their moments parallel (Fig. 2). The C—C—C angles are then increased to 180°; at the same time, the

---

[19] Wiswall and Smyth, *J. Chem. Physics*, 1941, 9, 352.

[20] Cartwright and Errera, *Proc. Roy. Soc.*, *A*, 1936, 154, 138.

H—C—H angles may well be reduced, thus making the methylene groups more polar. Because these group moments are in any case small, and because the restoring force constants are large, this process gives rise only to a very small polarisability. Moreover, the vibration frequency of such movements would be high, because the mass of a methylene group is small; so there would be no absorption of energy at dielectric frequencies due to them.

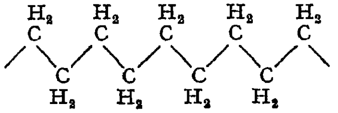


FIG. 1.

If the paraffin chain were supposed to assume an irregular, looped configuration (Fig. 3), as it is believed to do in the elastic hydrocarbons, it would still be non-polar. The loops might be expected to behave in some degree as polar units, because the methylene groups therein are to some extent aligned so that their moments give small local resultants, and they might orient as such in a field, with concurrent changes of the valency angles, thereby giving rise to atom polarisation. It is not at all obvious whether this atom polarisation would be different from that of the linear configuration, or what the difference might be; but what does seem probable is that the frequencies of vibration of the larger units would be lower than that of a methylene group, especially if braked by rubbing either against adjacent loops of the same molecule or against other molecules, with the result that a liquid or a non-oriented solid paraffin might show a small dielectric loss which a crystalline or an oriented solid paraffin would not.

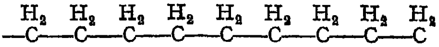


FIG. 2.

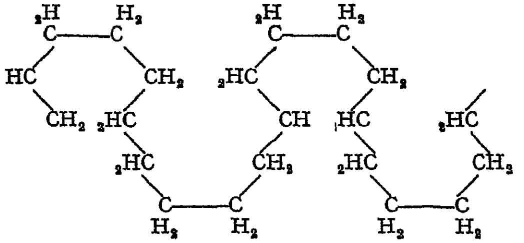


FIG. 3.

Orientation Polarisation.—By far the greatest part of the applications of dielectric measurements to chemical problems have been concerned with permanent electric dipole moments. They divide into two main groups: (i) studies of the moments of isolated molecules, (ii) studies of the effect on polar molecules of their surroundings and of their interactions with these, in liquid or solid solutions, or in the liquid or solid state of the pure substance.

The first group subdivides into (i) the use of the vector character of the electric moments of bonds or groups to decide questions of stereochemistry, and therefore to investigate factors influencing molecular geometry; (ii) the use of the magnitude of moments to investigate the nature of the electron distribution and so to examine the nature of bonds, *e.g.* to find whether they are mainly covalent or electrovalent or whether "oxy-bonds" are dative links or multiple links,[21] or to search for electron drifts consequent upon resonance.[22]

The second group subdivides into studies of (i) the effect of dipole-dipole and dipole-solvent interaction on the orientation polarisation, (ii) change of polarisation with change of phase, (iii) dielectric loss in the liquid and solid states, (iv) the effect of polarity upon thermodynamic properties.

[21] Phillips, Hunter and Sutton, *J. Chem. Soc.*, 1945, 146.

[22] Ref. [14], p. 69; Wheland, *The Theory of Resonance*, New York, 1944, Hannay and Smyth, *J. Amer. Chem. Soc.*, 1943, 65, 1931.

Our particular concern is with the possible improvement of these applications, or with the development of new ones. No essentially new methods of measuring moments have been developed recently; but, because of recent improvements in ultra-high-frequency technique, the dielectric loss method of determining them may be used more in future than it has been in the past. Nor, with one possible exception, are there in view any applications novel in method; though new chemical problems to which existing methods can be applied come in a steady succession. Three difficulties which remain imperfectly solved are (i) how to allow for the solvent effect, (ii) how to compute the moments induced by a dipole in its near surroundings, (iii) how to measure the moments of individual bonds, (iv) how to predict accurately and from first principles the magnitude of the moment of a bond or molecule (though, fortunately, it is possible to make rough predictions using parameters such as the electronegativities of elements). These limit the scope of such applications. The possible exception, noted above, is the direct determination of the variation of moment of a bond with internuclear distance, which could throw light on its nature.[23]

Of the applications in the second group, two may be singled out for mention. One is the use of polarisation measurements for detecting intermolecular interaction. This has already been tried, but the general difficulty is that there are two unknowns, *viz.* the degree of interaction and the polarity of the complex.[24] In one case this difficulty has been overcome by a special trick for determining the former independently;[25] but some more general method is needed. Partition methods seem possible, and perhaps infra-red absorption spectra may be applied as they already have been to hydrogen bonding, or dielectric loss measurements may be usable. The chemical application of these latter measurements is, in fact, the remaining topic.

In the earlier theories of the dielectric loss which arises when dipoles cease to rotate in condensed phases, a coefficient of viscosity was introduced; but this was *per se* of limited use to the chemist for interpreting the behaviour of liquids, because viscosity itself required to be related to energies and probabilities; while for solids it had little conceptual value at all. More recent theories attempt to relate dielectric loss to these quantities.[26] Chemists would like to learn from the present discussion how far these attempts may be considered satisfactory. Semi-empirical applications of loss measurements to investigating the strength of bonding and the conditions for rotation of molecules in different states of aggregation and in different surroundings are already extensive. If there is a sound and certain basis for interpreting the results in terms useful to the chemist, such investigations have great potentialities.

The Physical Chemistry Laboratory,
Oxford.

[23] Bell and Coop, *Trans. Faraday Soc.*, 1938, 34, 1209.
[24] Hammick, Norris and Sutton, *J. Chem. Soc.*, 1938, 1755.
[25] Fairbrother, *J. Chem. Soc.*, 1945, 503.
[26] Kauzmann, *Rev. Mod. Physics*, 1942, 14, 12.

MOLECULAR FREEDOM IN SOLID DIELECTRICS.

BY CHARLES P. SMYTH.

Received 28th February, 1946.

The dielectric properties of solid substances consisting of polar molecules depend largely upon the freedom of the molecules or of their polar parts to alter their orientation when the solid is placed in an electric field. It is the purpose of this paper to examine, in a qualitative fashion, the evidence given by experimental data as to the factors which determine the extent of this freedom and, hence, the dielectric behaviour of the material.

If the molecules of a substance are without dipole moment the liquid has a small dielectric constant, slightly larger than the square of the refractive index for visible light, and no dielectric absorption. The only change produced by solidification is the small change of dielectric constant produced by change of density, usually an increase. If the molecules possess dipole moments, the liquid has a larger dielectric constant, usually between 3 and 80, depending upon the size of the moment and the number of molecules per cm.^{3}. Dielectric absorption occurs in a frequency region depending approximately upon the size and shape of the molecules, the viscosity of the material and the temperature. The normal behaviour on solidification is a fixing of the molecules in a space lattice which prevents molecular orientation in an externally applied electric field and thus reduces the dielectric behaviour of the solid to that of a non-polar substance.

It is the not infrequent departure from normal behaviour that is of primary interest in the present discussion. It was found,[1, 2] in accordance with a prediction [3] based on the behaviour of the specific heat, that the dielectric constant of the polar substance, hydrogen chloride, continued to rise with decreasing temperature in the solid state as in the liquid until the occurrence of a phase transition, which apparently fixed the molecules to give a sharp drop in dielectric constant from 17·3 to 4. Instead of dropping sharply in the usual manner of polar substances on freezing, the dielectric constant rose slightly because of the greater density of the solid. More or less similar behaviour was observed for hydrogen bromide,[2] hydrogen iodide,[2] hydrogen sulphide,[4] hydrogen selenide,[5] arsine [5] and other small molecules or groups [6] including water as ice,[6, 7] and as water of hydration in salts,[8, 9] and for a few large molecules,[10, 11] such as camphor and borneol. The transition from high dielectric constant to low was found commonly to be sharp, sometimes involving some hysteresis, but not infrequently gradual and dependent upon the frequency of the field used in the measurement. It was evident that, in these solids above the transition points, the molecules possessed freedom of orientation like that in the liquid state, a freedom described as rotational freedom, although it, probably, meant the possession of sufficient rotational energy to permit of sufficiently frequent passage over the potential energy barriers between different preferred orientations to give an essentially random orientation. Rotational freedom such as this would give the molecules spherical symmetry in the lattices of the solid and result in lattices corresponding to the arrangements of close-packed spheres. Of the first eighteen substances of known crystal form which were found to show rotational freedom of the molecules, it was found [12] that twelve belonged to the cubic system in the region of rotation, four to the hexagonal, one, *d*-camphor to the cubic below a transition and to the hexagonal above it, and one, hydrogen iodide, to the tetragonal. The loss of rotational freedom commonly resulted in the change of the solid from an isotropic to an

---

[1] Cone, Denison and Kemp, *J. Amer. Chem. Soc.*, 1931, **53**, 1278.
[2] Smyth and Hitchcock, *ibid.*, 1933, **55**, 1830.
[3] Pauling, *Physic. Rev.*, 1930, **36**, 430.
[4] Smyth and Hitchcock, *J. Amer. Chem. Soc.*, 1934, **56**, 1084.
[5] Smyth and McNeight, *ibid.*, 1936, **58**, 1723.
[6] Smyth and Hitchcock, *ibid.*, 1932, **54**, 4631.
[7] Murphy, *Trans. Amer. Electrochem. Soc.*, 1934, **65**, 133.
[8] Errera and Brasseur, *Physik. Z.*, 1933, **34**, 368.
[9] Errera and Sack, *Trans. Faraday Soc.*, 1934, **30**, 687.
[10] Yager and Morgan, *J. Amer. Chem. Soc.*, 1935, **57**, 2071.
[11] White and Morgan, *ibid.*, 1935, **57**, 2078.
[12] Smyth, *Chem. Rev.*, 1936, **19**, 329.

anisotropic lattice, which made possible investigation of the so-called rotational transitions in crystals by means of a polarising microscope.[13] The complexity of the relations in the solid has thus far prevented successful quantitative treatment[14, 15] of the phenomenon as a whole, although progress was made toward an approximate representation of the effect of hindered molecular rotation in the hydrogen halides,[16] and the dielectric behaviour of hexasubstituted benzenes in the solid state could be approximated[17] in terms of a model in which rotational motion could occur from one to another of two equilibrium orientations of different potential energies, with the frequency and temperature range of anomalous dispersion determined by the height of the intervening potential barrier.[18]

In view of the difficulties in the way of a quantitative theoretical treatment of the general problem of molecular rotation, further examination of the now considerable body of experimental data is desirable. Dielectric evidence of rotation was found for three derivatives of ethane,[19, 20] for derivatives of cyclopentane and cyclohexane,[11, 20] and for a large number of derivatives of camphane.[10, 20] In all these cases, as in those of the small molecules previously mentioned, the molecules were not far removed from sphericity in form, so that the potential barriers to rotational motion were sufficiently low to permit of their being frequently crossed. Twelve penta- and hexasubstituted benzenes,[17] in which the substituent groups were small and about equal in size, showed rotation, presumably, in the plane of the ring. The attachment of a group which protruded much beyond the others either in the plane of the ring or perpendicular to it blocked rotation. The less symmetrical of the molecules of these substituted benzenes showing rotation appeared to cease abruptly their rotations at sharp transitions like that observed for hydrogen chloride, while the more symmetrical tended to show a gradual cessation of rotation over a range of falling temperature, giving anomalous dispersion and absorption like that in a viscous liquid. Ethyl stearate and cetyl alcohol were found to show apparent rotational transitions above which the orientational freedom of the dipoles was attributed to rotation around the long axes of the rod-shaped molecules,[21] while evidence indicating the rotation of long carbon chains in crystals was obtained from X-ray measurements.[18]

Tetrasubstituted methanes in which the substituents were halogens and methyl groups showed, in the solid state, for some distance below the freezing points, freedom of orientation of the molecules like that in the liquid.[13, 22] The orientational freedom of these molecules is due to the symmetry of their form, which is approximately that of a tetrahedron with its corners so rounded as not to differ greatly from a sphere. When one of the methyl groups of *t*-butyl chloride is replaced by an ethyl to form *t*-amyl chloride, the protrusion of the ethyl group causes sufficient interlocking of the molecules in the lattice to prevent rotation.[23] The similar change from methylpentachlorobenzene to ethylpentachlorobenzene blocks rotation in the latter.[17] In the hexasubstituted methylchlorobenzenes, successive replacement of methyl by chlorine, which should increase the local molecular field and, therefore, the potential barriers, normally raises the temperature at which molecular rotation sets in.[17] In the methylchloromethanes,[22] the change from neopentane to *t*-butyl chloride raises the rotational transition temperature 79°, but the transition of the dimethyldichloro compound is 31° lower than that of *t*-butyl chloride, while methyl chloroform and carbon tetrachloride are 5 and 6° higher respectively. In the anisotropic lattices in which these compounds exist below their transition points, differences in the packing of the molecules can easily account for the irregular variation of the transition temperatures. Once the molecules are rotating, or passing frequently over the potential barriers hindering rotation, their almost spherical symmetries give rise to isotropic lattices, which melt within a range of only 13·7°. In the cases of the hydrogen halides and the *t*-butyl halides, increase in the size of the halogen atom from chlorine to bromine to iodine lowers the temperature necessary to give the molecules rotational freedom in the crystal. The replacement of the halogen of a *t*-butyl halide by the nitro group, which differs from a methyl group in shape more than does a halogen and gives a much larger dipole moment, still does not prevent orientational freedom in the solid,[24] and 1, 2, 3, 4-tetramethyl-5, 6-dinitrobenzene and 5-nitro-3, 4-dichloro-*o*-xylene also show rotational freeeom.[17]

---

[13] Baker and Smyth, *J. Amer. Chem. Soc.*, 1939, **61**, 2798.

[14] Frenkel, Todes and Ismailow, *Acta Phys.-Chim. U.R.S.S.*, 1934, **1**, 97.

[15] Fowler, *Proc. Roy. Soc., A.*, 1935, **149**, 1; *Statistical Mechanics*, Cambridge, 1936, 2nd edition, chap. xxi.

[16] Kirkwood, *J. Chem. Physics*, 1940, **8**, 205.

[17] White, Biggs and Morgan, *J. Amer. Chem. Soc.*, 1940, **62**, 16.

[18] White, *J. Chem. Physics*, 1939, **7**, 58.

[19] White and Morgan, *ibid.*, 1937, **5**, 655.

[20] White and Bishop, *J. Amer. Chem. Soc.*, 1940, **62**, 8.

[21] Baker and Smyth, *ibid.*, 1938, **60**, 1229.

[22] Turkevich and Smyth, *ibid.*, 1940, **62**, 2468.

[23] Turkevich and Smyth, *J. Amer. Chem. Soc.*, 1942, **64**, 737.

In order to investigate the effect of molecular environment upon orientational freedom, dielectric constant measurements were made [25] upon mixtures of *t*-butyl chloride and carbon tetrachloride, which had been found [22] to form a continuous series of solid solutions. The dielectric measurements made possible the construction of a much more complete equilibrium diagram for the binary system than could be drawn from thermal measurements. The addition of one component to the other lowers the freezing point somewhat, but lowers the temperature of the principal rotational transition more, thus increasing the temperature range within which the molecules possess rotational freedom in the solid. The phase changes involved do not appear abnormal for a binary system. Examination with a polarising microscope showed that the lattices of the solid solutions were isotropic as long as the molecules possessed rotational freedom comparable to that in the liquid. With the disappearance of rotational freedom, the lattices became anisotropic.

The crystalline solids, in which molecular rotational freedom has been observed, are usually formed from the liquid with little or no supercooling. Freezing usually brings about a small break in the dielectric constant-temperature curve because of change of density. Moderately unsymmetrical molecules, like the somewhat pear-shaped *i*-butyl and *i*-amyl bromide molecules and similar substituted or branched hydrocarbons, can pack randomly in the liquid with an economical use of volume. Consequently, although the molecular arrangement is, on the whole, a disordered one, the efficient packing may allow persistence of the liquid arrangement far below the normal crystallisation temperature and thus give rise to supercooling and glass formation.[26] The dielectric behaviour of such glasses is that of a viscous liquid. On cooling through the temperature of the melting-point, the dielectric constant curves for *i*-butyl bromide,[27] *i*-amyl bromide,[27] and *i*-butyl chloride [28] showed no discontinuity at the temperature of the melting-point, but strong anomalous dispersion at much lower temperatures with dropping of the dielectric constant almost to 2·5 at kilocycle frequencies. These glasses were mechanically rigid at temperatures well above the region of anomalous dispersion. The two bromide glasses crystallised slowly at liquid air temperature so that only slight anomalous dispersion was observed on warming and the dielectric constant rose sharply at the melting-point. The *i*-butyl chloride remained a glass at liquid air temperature for periods up to 100 hours and, on warming, crystallised completely in the region of anomalous dispersion with the dielectric constant falling to the value for rigidly fixed molecules and rising sharply at the melting-point to that of the liquid. 1, 2-dichloroisobutane[28] formed a glass which did not crystallise, but, otherwise, resembled the other glasses in its dielectric behaviour. In the temperature region above that of anomalous dispersion for kilocycle frequencies, the dielectric behaviour of these glasses was indistinguishable from that of a liquid or of a crystalline solid in which the molecules possessed rotational freedom.

---

[24] Unpublished measurements by Dr. W. P. Conner and Dr. A. Turkevich in this laboratory.

[25] Conner and Smyth, *ibid.*, 1941, **63**, 3424.

[26] Baker and Smyth, *Ann. New York Acad. Sci.*, 1940, **40**, 447.

[27] Baker and Smyth, *J. Amer. Chem. Soc.*, 1939, **61**, 2063.

[28] Turkevich and Smyth, *ibid.*, 1942, **64**, 737.

Although the change of dielectric constant produced by freezing or melting is, for the most part, an exceedingly sharp one, the dielectric constant of the solid just before melting or just after freezing is higher than it is at temperatures well below the melting-point and decreases at a decreasing rate with some anomalous dispersion as the temperature decreases below the freezing-point. A portion of this effect, at least in many cases, is attributable to the presence of impurities, which may spread the last part of the process of solidification or the first part of the process of fusion over a considerable range of temperature with the possibility of a small Maxwell-Wagner effect to give a slight dielectric absorption. However, the effect appears to be of too general occurrence and, often, of too large size to be generally explicable in terms of impurities. Specific heat measurements near the melting-point give evidence of a similar loosening of the solid structure, which is often referred to as "premelting." This premelting was evident, for example, to a small extent, in carefully purified nitrobenzene,[6] over a temperature range of nearly 40° in isopropyl bromide,[29] where impurities may have contributed somewhat to the effect, and in ethyl undecylate,[31] where the dielectric constant increased at a uniform rate for about 25° from the lowest measurement at —58·8° C. and, then at an increasing rate as the temperature of melting, —22·5°, was approached. A similar loosening of structure is often apparent in the temperature region below a sharp rotational transition and may be termed "prerotation." The effects of prerotation and premelting upon dielectric constant are small in comparison with that of rotational freedom of the molecules as in a liquid. They cause a usually small rise in a low dielectric constant with rising temperature as loosening of the lattice structure permits increase in molecular orientation, while rotational freedom of the molecules as in a liquid causes a large decrease with rising temperature in a relatively high dielectric constant, the polarisation varying inversely as the absolute temperature.[2]

It has been pointed out that dipole orientation can sometimes occur around the long axis of a rod-shaped molecule. The very large molecules of the polymers of *ω*-hydroxydecanoic acid in solutions at frequencies too high to permit of any molecular orientation[30, 31] show dipole moments corresponding to freedom of orientation, within the molecules, of the group dipoles, which are separated from one another by chains of nine carbon atoms. In the chains of solid polymers, now one, now another segment of a chain may give dipole orientation and these segments may vary in length between the two improbable extremes of a single monomeric unit and the whole extended chain.[32] The behaviour of the dielectric constant and loss of a high polymer, which has been successfully represented in such terms,[32] is thus similar to that of a viscous liquid or a glass composed of small molecules. Indeed, a transition temperature has been found[33] in polyvinyl chloride at which a process describable as an internal melting occurs, a phenomenon analogous to the setting-in of molecular rotation in a crystalline solid.

---

[29] Baker and Smyth, *J. Amer. Chem, Soc.*, 1939, **61**, 1695.

[30] Wyman, *ibid.*, 1938, **60**, 328.

[31] Bridgman, *ibid.*, 1938, **60**, 530.

[32] Fuoss and Kirkwood, *ibid.*, 1941, **63**, 385.

[33] Fuoss, *ibid.*, 1941, **63**, 369.

Some crystalline solids, glasses, viscous liquids and solid polymers may, evidently, show similar dielectric behaviour as the result of dipole orientation, the restriction of which is dependent upon molecular shape, packing in the lattice, and internal energy. The heats of transition and of fusion show that a solid phase below a rotational transition may differ more in energy from the phase above the transition than this rotating phase differs from the liquid.[26, 29, 31] Freedom of molecular rotation as evidenced by a dielectric constant value much higher than the square of the refractive index implies essentially the possession of sufficient molecular rotational energy to permit of frequent passage over restricting potential barriers. Intramolecular rotation or cumulative twisting about bonds may give dipole orientation in large molecules. Premelting and prerotation would seem to be the effects of the orientation of molecules here and there in the solid. As the effects increase in number, they co-operate to produce a rotational transition or fusion as the case may be.

Frick Chemical Laboratory,
Princeton University,
Princeton, N.J.,
U.S.A.

---

[31] Smyth, *Proc. Amer. Phil. Soc.*, 1936, 76, 485.

THE DIELECTRIC RELAXATION OF MIXTURES OF DIPOLAR LIQUIDS.

By Adolf Schallamach.

Received 4th February, 1946.

The dielectric behaviour of dipolar liquids in non-polar solvents has received a great deal of attention in the literature but no reference to work on mixtures of dipolar liquids appears to have been published. The only paper bearing on a similar subject which has come to the author's notice is by Alexandrov and Dzhian [1] who have reported that vulcanised rubber swollen in dipolar solvents give loss factor curves with only one maximum. From this result it appeared that it might be of interest to study the dielectric relaxation of *mixtures* of dipolar liquids because it should be borne in mind that one peak of the loss factor curve must be taken to signify the prevalence of one relaxation mechanism. For the definition of the dielectric constants and concepts we refer to an earlier publication.[2]

Materials and Experimental Procedure.

The choice of suitable compounds is somewhat restricted for the following reasons. The dielectric dispersion ranges of the pure liquids should be sufficiently separated so that their loss factor maxima do not merge in the mixture. This condition involves generally quite different molecular weights and the dispersion range of the lighter component will, even at high radio frequencies, occur at rather low temperatures. The liquids used must therefore not crystallise or it must be possible easily to supercool them. Finally, they must be mutually miscible. Unless recourse is had to difficult organic syntheses, the number of suitable compounds is limited.

---

[1] Alexandrov and Dzhian, *Tech. Phys. U.S.S.R.*, 1938, 5, 836.

[2] Schallamach, *Trans. Faraday Soc.*, 1946, 42, 495.

The liquids used in this investigation were: geraniol, di-dihydrocitronellyl ether (whose dielectric properties have already been reported [2]), and *n*-propyl alcohol, glycerol, di-*n*-butyl ether, *iso*-amyl bromide, citral and geranic acid, which were redistilled commercial products.

The experimental procedure was as described before [2] but a specimen condenser with an air capacity of only 1 *p*F was used for liquids with higher dielectric constants.

Results.

Fig. 1 gives the results obtained with an equimolecular mixture of geraniol and *n*-propyl alcohol at three different frequencies. The loss factor curves (ϵ'') have only one maximum. At 134 Kc./s., the ϵ''-curve rises again at higher temperatures, owing to the onset of electrolytic conduction. The broadening of the curves at higher temperatures and frequencies is a commonly observed phenomenon and largely due to the approximately exponential temperature dependence of the relaxation time

$$\tau = \tau_0 \, e^{E/RT}$$

which makes it obvious that the widths of the curves should be compared in terms of the reciprocal absolute temperature.

The maxima of the loss factor curves of the pure liquids occur at the temperatures given below, and it will be seen in Fig. 1 that the curve is sharpest where the separation of the individual maxima is largest. Incidentally, the curious fact should be noted that above room temperature the relaxation time of *n*-propyl alcohol is actually larger than that of geraniol.

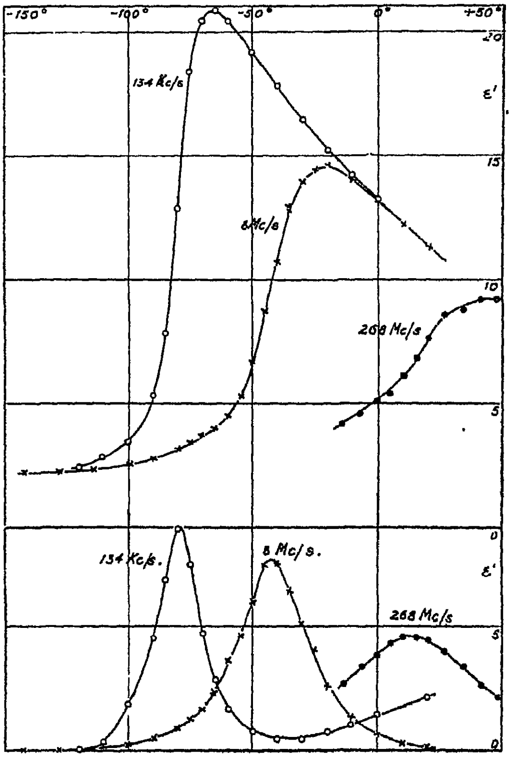


FIG. 1.—Dielectric constants of an equimolecular mixture of geraniol and *n*-propyl alcohol at 3 frequencies vs. temperature °C.

	134 Kc./s.	8 Mc./s.	268 Mc./s.
n-propyl alcohol [3] . .	— 112° C.	— 62·5° C.	30° C.
Geraniol [2]	— 61·5° C.	— 28·5° C.	16° C.

---

[3] Kobeko, Kuvshinsky and Shishkin, *Tech. Phys. U.S.S.R.*, 1938, 5, 413.

The curves for a mixture of 2/3 by volume of *n*-propyl alcohol and 1/3 glycerol at 9 Mc./s. are shown in Fig. 2, and here again the loss factor curve has only one maximum. The pure liquids have their maxima at:

n-propyl alcohol[1] . — 60·5° C.
Glycerol[3] . . — 10° C.

We have dealt so far with highly associated liquids which may be expected to form complex units. Figs. 3 and 4 show, however, that mixtures of liquids not conspicuous for their association exhibit only one loss factor maximum. In particular, the ethers of Fig. 3 are as little associated as dipolar compounds can be expected to be, but it should be pointed out that the curves of these mixtures are somewhat broad.

The comparatively large dielectric constants of citral shown in Fig. 4 may be attributed to conjugation between the double-bond and the carbonyl group. The curves for pure *iso*-amyl bromide have been omitted in this figure, to avoid confusion, but are reproduced in Fig. 5. The value of the static dielectric constant found for this substance is slightly larger than that given by Baker and Smyth,[4] and we have not observed hysteresis phenomena such as have been reported by these authors.

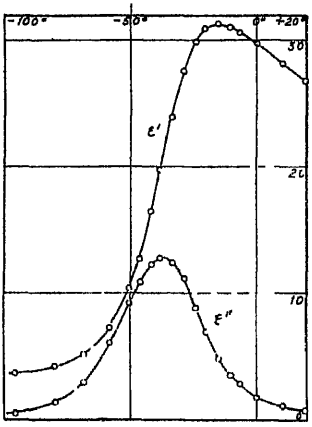


FIG. 2.—Dielectric constants of a mixture containing 2/3 by volume of *n*-propyl alcohol and 1/3 of glycerol at 9 Mc./s. vs. temperature °C.

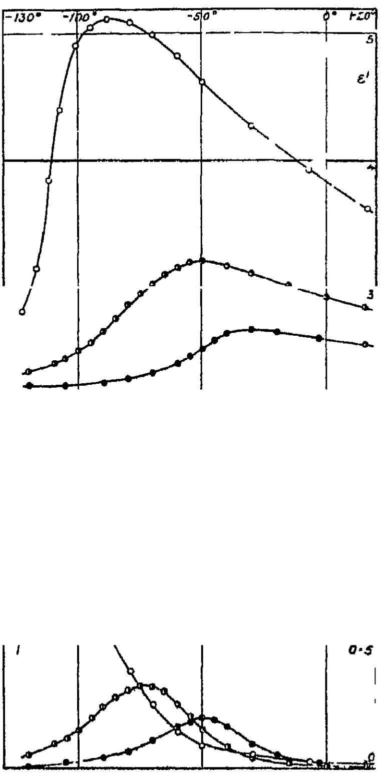


FIG. 3.—Dielectric constants of di-*n*-butyl ether o, di-dihydrocitronellyl ether ●, and of an equimolecular mixture of these 2 ethers ◐, at 20 Mc./s. vs. temperature °C.

The mixture of *iso*-amyl bromide and geraniol (Fig. 5) presents a behaviour different from the mixtures discussed above. As seen from the two broad but definite maxima and from the shape of the ε' curve, this mixture does not appear to have one single relaxation mechanism.

Mixing geraniol with di-*n*-butyl ether has not so marked an effect (Fig. 6) most probably because the dielectric constants of the ether are small. Its dielectrically separate presence may be observed on the low-temperature side of the maximum where the loss factor declines only very slowly.

---

[4] Baker and Smyth, *J. Am. Chem. Soc.*, 1939, 61, 2063.

As it appeared from these results that the mixture of an associated and a non-associated liquid gave two maxima, a mixture of *iso*-amyl bromide and *n*-propyl alcohol was investigated, as association is stronger in the latter than in geraniol. Fig. 7 shows indeed that the separation of the maxima is here very pronounced. The curve for ε′, too, gives definite evidence of two different mechanisms as it falls in two steps and is even seen to rise slightly where the loss factor has a minimum. Fig. 7 would almost suggest that the mixture had separated into two phases but visual observation did not indicate any macroscopic inhomogeneity.

It should be noted that the maxima of pure *n*-propyl alcohol and citral occur at similar temperatures and so Figs. 4 and 7 are particularly appropriate for comparison of the two types of curves encountered in these mixtures.

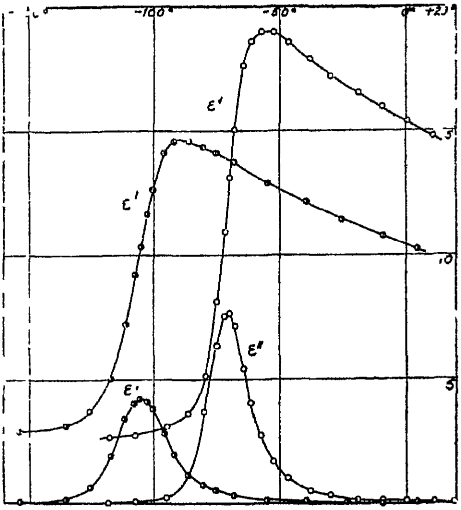


Fig. 4.—Dielectric constants of citral o and of a mixture containing 73·5 mole % of iso amyl bromide and 26·5 mole % of citral ◑ at 9 Mc./s. vs. temperature °C.

Geranic acid (Fig. 8) was chosen for the study of another type of associated liquid. The association of acids is well known to be different from that of alcohols. Dielectrically, its effect in acids is to make the dielectric constant smaller than would be expected from the dipole moment [5] whereas in alcohols it is larger than the normal value. At 20° C. the static dielectric constant of geraniol is 6·8 and of geranic acid 4·9, in keeping with this general experience.

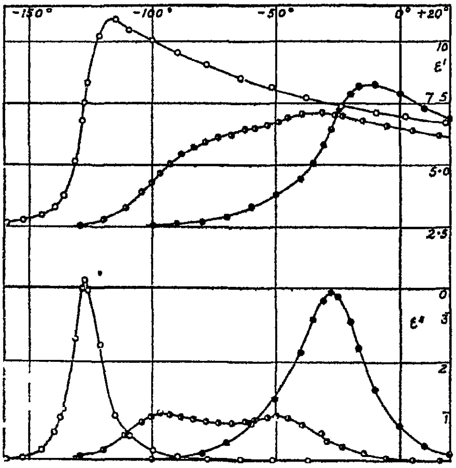


Fig. 5.—Dielectric constants of *iso*-amyl bromide o, of geraniol ● and of an equimolecular mixture of both compounds ◑ at 9 Mc./s. vs. temperature °C.

Measurements on mixtures of geranic acid and both di-*n*-butyl ether and *iso*-amyl bromide at 9 Mc./s. gave inconclusive results which just indicated that the dielectric curves had not the normal shape. The measurements reproduced in Fig. 8 were therefore carried out at a lower frequency, and consequently lower temperatures, where the separation of the individual maxima is greater. It will be seen that the loss factor of the mixture gives a broad, flat-topped curve and, although ε′ of the mixture is considerably higher than that of the pure acid, the absolute values of ϵ'' at the maxima are practically equal. The dielectric behaviour of this mixture is different from that shown by the corresponding geraniol mixture, with which it might be compared, but there is sufficient evidence to show that the relaxation mechanism is not sharply defined. This mixture appears to be intermediate in properties between those described in Fig. 4 and Fig. 7.

---

[5] Kumler, *J. Am. Chem. Soc.*, 1935, 57, 600.

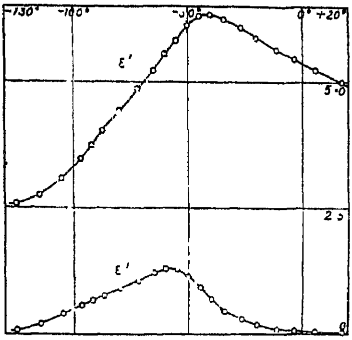


FIG. 6.—Dielectric constants of an equimolecular mixture of di-n-butyl ether and geraniol at 9 Mc./s. vs. temperature °C.

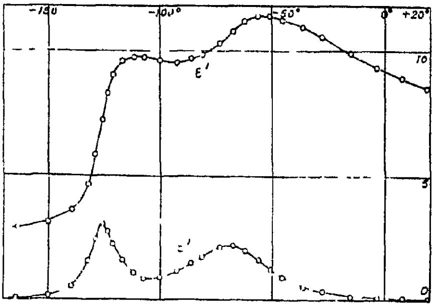


FIG. 7.—Dielectric constants of a mixture containing 23·5 mole % of n-propyl alcohol and 76·5 mole % of *iso*-amyl bromide at 9 Mc./s. vs. temperature °C.

Discussion.

The results presented above may be summarised by saying that, of the liquids studied, mixtures containing only associated or only non-associated liquids have normal though sometimes broad loss factor curves with one definite maximum, but that the loss factor curves of mixtures of associated and non-associated liquids are not normal and have sometimes two definite maxima. It must be added, however, that the appearance of two maxima depends partially on accidental factors such as the relative concentrations. It has, for example, been found that a mixture *iso*-amyl bromide-geraniol containing only 26 mole % of the alcohol gives a loss factor curve similar to Fig. 6.

The published theories of dielectric relaxation do not deal specifically with mixtures of dipolar liquids but it would appear from the prevalent ideas on this subject that any binary mixtures should have two dispersion ranges because, roughly speaking, the hindering forces envisaged by Debye as well as the potential troughs of the so-called kinetic theory depend on the molecule and so give a separate relaxation time for each molecular species present.

Judged from the experimental evidence, however, such behaviour is the exception rather than the rule and is confined to special mixtures, *i.e.* those containing associated and non-associated liquids. Taking into account Alexandrov and Dzhian's as well as our results we feel inclined to think that normally the dielectric relaxation of binary mixtures is dominated by only one mechanism.

Alexandrov and Dzhian proposed that complex polar groups are formed in the rubber vulcanizate which groups move as one in dielectric relaxation. We find it difficult to visualise such complexes as dielectric units in a liquid, particularly as then one would expect much longer relaxation times and higher activation energies than are commonly found. Without going into any detail the following view is suggested. The elementary process in dielectric relaxation cannot be directly connected with individual molecules but is a disturbance of an appreciable region in the liquid, involving a volume nearly large enough to be representative in its composition of the bulk concentrations. On this basis any liquid, even if a mixture, would have only one relaxation time.

It would follow that the mixture of an alcohol and a non-associated liquid is not microscopically homogeneous and that the associative forces are strong enough to maintain in the mixture regions containing mostly alcohol molecules. It is well known that such mixtures show large positive departures from Raoult's law, and in extreme cases have limited miscibility.[6]

A clear-cut division of all mixtures into two classes cannot be expected and a certain degree of association is well known to exist in all dipolar liquids. This consideration makes it understandable why, for example, the loss factor curve of the mixture *iso*-amyl bromide-citral is relatively broad. As for the acids, there is evidence that association leads here mostly to dimers and so its influence is spatially restricted. This would account for the difference between the geraniol and geranic acid mixtures.

There is no reason to assume that dielectric relaxation is fundamentally different in mixtures and in pure liquids. If the above proposition be accepted it follows that this physical phenomenon quite generally involves a relatively large volume in the dielectric. Fröhlich and Sack [7] make assumptions which appear somewhat similar but our views seem to be opposed to those of Kauzmann.[8]

Finally, certain conclusions ought to be drawn with respect to other physical processes which have been connected with dielectric relaxation, particularly viscosity. It has been shown [2, 3, 8] that in some cases there is strong reason to believe that the elementary processes leading to these two phenomena are identical. In a previous paper on this subject we shortly discussed the difficulties encountered with simple molecular models of viscous flow of which one was the necessity of co-ordination between the movements of different units. If, by analogy with dielectric relaxation, it is assumed that comparatively large volumes are disturbed before, or when, viscous flow takes place it is easier to conceive the possibility of co-ordination within these regions than if only the movement of individual molecules is taken into consideration.

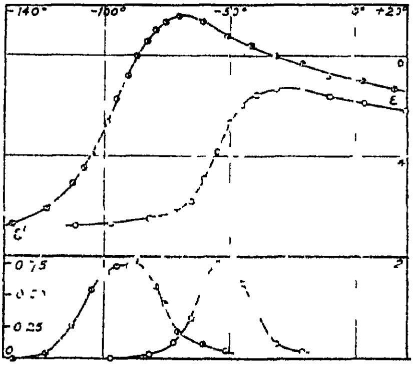


Fig. 8.—Dielectric constants of geranic acid o and of a mixture containing 55 mole % of geranic acid and 45 mole % of *iso*-amyl bromide ◑ at 325 Kc./s. vs. temperature °C.

Summary.

It has been found experimentally that the curves giving the dielectric constants of a number of binary mixtures as functions of the temperature are normal if the mixtures contain only associated or only non-associated liquids but are not normal for mixtures of associated and non-associated liquids. In the latter case, the loss factor curves exhibit under certain circumstances two maxima. It is suggested that dielectric relaxation involves the disturbance of relatively large regions in the liquid and that a mixture of associated and non-associated liquids is not microscopically homogeneous.

---

[6] Hildebrand, *Solubility*, New York, 1936.

[7] Fröhlich and Sack, *Proc. Roy. Soc., A*, 1944, **182**, 388.

[8] Kauzmann, *Rev. Mod. Physics*, 1942, 14, 1.

This work forms part of the programme of fundamental research undertaken by the Board of the British Rubber Producers' Research Association.

48 *Tewin Road,*
Welwyn Garden City, Herts.

THE RADIO-FREQUENCY ABSORPTION SPECTRA OF SOLUTIONS OF ELECTROLYTES.

By Jan Forman and Dennis J. Crisp.

(*Communicated by* Professor E. K. Rideal.)

Received 8*th May*, 1946.

It is a well-known fact that three types of polarisation must be included in the general dielectric theory: (1) electronic, (2) atomic and (3) molecular. It is the molecular, or dipole, polarisation, due to the orientation of polar molecules or groups attached to molecules, which accounts for the greater part of the dielectric constant of polar molecules, and for most of the power loss or absorption when these molecules are subjected to alternating electric fields of frequencies up to 10″ c.p.s.

The maximum loss occurs at a frequency (f), of pulsatance (ω) $= 2\pi f$, when $\omega\tau = 1$, τ being referred to as the relaxation time of the molecule or polar group, and being a function of the molecular dimensions, the viscosity, and the absolute temperature. This loss is due to the lag of the orientation setting up an in-phase component of the displacement current and thereby causing ohmic loss in the dielectric with consequent rise of temperature. In the frequency region corresponding to the maximum loss the phenomenon of anomalous dispersion of the dielectric constant is also observed.

In addition to the absorption of energy as a result of electronic, atomic and molecular displacement, a further type of absorption is predictable in the case of electrolytes.

In an ionic crystal a characteristic frequency exists due to the forces holding the ions in the lattice, and when solution occurs a loose quasi-crystalline structure remains. According to the Debye-Falkenhagen theory,[1] this structure is caused by the tendency of each positive ion to be surrounded by an " atmosphere " of negative ions, and vice versa.

The quasi-crystalline structure of electrolytic solutions is opposed by thermal motion of the ions, and it can also be displaced under the influence of a static electric field, which will account in part for the anomalous behaviour of strong electrolytes. In an alternating field, of frequency determined by the relaxation time of the ionic " atmosphere," a maximum energy absorption will occur in a manner analogous to the absorption in di-polar substances. Thus, solutions of electrolytes exhibit a fourth type of polarisation, which may be termed " ionic polarisation " having a relaxation time τ with maximum power loss at a pulsatance ω, satisfying the relation $\omega\tau = 1$.

Debye and Falkenhagen[2] have shown that this relaxation time, Θ, is given by

$$\Theta = \frac{8{\cdot}85 \times 10^{-11} \times D_0}{\Lambda_\infty \times \gamma^*}, \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where D_0 = dielectric constant of water (80),
Λ_∞ = equivalent conductivity of the solution,
γ^* = concentration of the solution (g. eq. per litre).

---

[1] Falkenhagen, *Electrolytes*, pp. 100, 176 *et seq.*, Clarendon Press.

[2] *Ibid.*, pp. 213 *et seq.*

The predicted relaxation times in a number of simple electrolytic aqueous solutions, have been found; the frequency corresponding to the maximum power loss is shown to be a function of the concentration of the solution and the ionic constants of the electrolyte, a convenient empirical relation being evaluated:

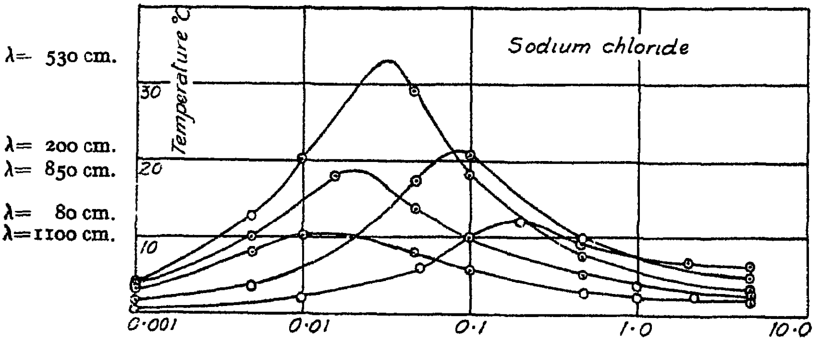
$$\lambda\gamma = K_e \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where λ = wavelength corresponding to frequency (f),
γ = concentration of the solution,
K_e = particular constant for each simple electrolyte.

Experimental.

Several binary electrolytes in aqueous solutions varying in concentration from 10 g. mole per l. to 0·00005 g. mole per l. have been examined at frequencies between 375 megacycles per second and 25 megacycles per second.

Two methods of measurement were employed; one being the determination of the temperature rise of a solution held in a quartz calorimeter within the condenser field of a high-frequency oscillator, and the other being the determination of the width of the resonance curve of a resonant line or circuit coupled to a high-frequency oscillator.



NaCl soln.—g. mole/l. = γ.

FIG. 1.

In the first method it was found impossible to make reliable measurements at one particular concentration of solution with several different frequencies, owing to variations in the power output of the oscillator when the frequency of oscillation was varied. Therefore the frequency and power output of the oscillator were held constant and the concentrations of the solution were varied. The temperature rise for a constant time of absorption of power at a constant input wattage was determined and a family of curves obtained, each one referring to a particular frequency, or wavelength (Fig. 1). These curves are fairly symmetrical and have a half-width of about 1·5 decades; the values of the maxima have no significance in relation to the whole family of curves since the power output of the oscillator varied considerably over the frequency range employed, but the positions of the maxima enable the concentrations to be determined at which the solutions exhibit a relaxation time corresponding to the pulsatance of the frequency referred to.

Owing to the very high power of the oscillator (about 300 watts c.w.), technical difficulties were reduced to a minimum. In practice 5 c.c. of the solution were placed in a small cylindrical quartz calorimeter, which was placed between two semi-circular brass plates, constituting a condenser, and attached near to the point of voltage antinode on a linear oscillating circuit. In this manner a small volume of the test solution is situated in the centre of a condenser field and enclosed by a vessel, the walls of which do not appreciably heat up due to power absorption from the high-frequency field. The initial temperature of the solution, well stirred, is noted by means of a mercury thermometer. The oscillator is then made to operate for a given time, generally of the order of 20 seconds, after which the solution is again stirred and the temperature noted. The whole operation is repeated for a series of solution concentrations at each frequency setting of the oscillator.

It can be seen, therefore, that although the peak values at various frequencies are different, this is entirely due to the fact that the output of the oscillator is not the same for all frequencies; nevertheless the significant information obtained from the experimental data is given by the positions of the maxima and the shapes of the curves.

The oscillator, which was powered by a special type of magnetron capable of delivering, if required, 600 watts c.w. between 70 cm. and 1200 cm. wavelength, was stabilised with regards to frequency by bridging the oscillator line circuit with an electrolytic impedance in place of the conventional, purely resistive metal bridge. The ionic relaxation time of the electrolytic aqueous solution, potassium dihydrogen phosphate, can be adjusted to the desired frequency so that it acts as a resistive line termination at only one particular small band of frequencies. By this means unwanted oscillations of other modes are eliminated, and a self-compensating load is applied to the oscillator. In practice it has been found that an increase of 100 % of either the d.c. applied to the anode of the oscillator or the d.c. applied to the electromagnets of the magnetron caused less than the minimum observable change of wavelength, *i.e.* 0·01 cm. at 80·00 cm. λ. The filament supply to the oscillator was stabilised by means of a saturated-reactor type transformer and although the anode H.T. supply was unstabilised, the magnetron was operated on a flat portion of its characteristic so that changes in amplitude became negligible.

Wavelengths of the oscillator up to two meters were measured on a pair of lecher wires in conjunction with a crystal detector. From this point to twelve meters the wavelengths were measured by an accurate wavemeter of Air Ministry type W1270.. By using harmonics of this wavemeter a cross-check could also be made on the accuracy of the lecher wire measurements down to a wavelength of 70 cm.

Experiments were carried out to determine the effect of the test solutions on the frequency of the oscillator. A solution of $LaCl_3$ 0·0278 N. was placed in the calorimeter and the oscillator adjusted to a frequency in the region of resonance. Radio-frequency power was applied in short bursts, with brief intervals in between during which the rise of temperature was noted. The results are shown in Table I. This represents a change of wavelength from 522·4 cm. to 522·6 cm. and is negligible, being less than 0·1 %.

Further experiments showed that the accuracy of the quantity of solution placed in the calorimeter was not an important factor, since a change of 0·1 c.c. only changed the oscillator frequency by 0·02 Mc./sec. These results are set out in Table II.

Further experiments showed that the position of the calorimeter in the condenser field was also of negligible importance, there being no detectable change of frequency for a shift of 8 mm. about a mean vertical position, and only a change of about 0·7 % of wavelength for the rotation of the slightly non-symmetrical calorimeter between the condenser plates, when it was found that the wavelength of the oscillator was 522·0 cm. with the shorter axis between the plates and 518·3 cm. with the longer axis between them.

Further experiment showed that the change of frequency due to the change of dielectric constant of the solution as the position of maximum loss was passed through was of the order of 0·2 % and therefore also negligible. These results are tabulated in Table III and shown in Fig. 2.

The oscillator is thus a relatively stable instrument.

The second method makes use of the determination of the shape of resonance curves obtained from a resonator tuned through the frequency of a stabilised oscillator. The stabilised magnetron oscillator injects a high-frequency wave of constant amplitude and of fixed frequency into a totally screened cage containing a linear resonator. This resonator consists of two parallel, Ag-plated, Cu rods, having a small, micrometer-adjusted condenser at the central position, and two, sliding Ag-plated Cu shields, which act as a resistive termination at each end of the line. These shields, or bridges are interconnected mechanically and move in a reciprocal fashion, both of them shortening or lengthening the line length simultaneously The advantage of this system is that the central condenser, in which the substance to be tested is placed, is always at a voltage antinode and therefore the field distribution across the substance is always similar. Such a resonator is normally used with non-conducting dielectric substances in the test condenser, the shift of the line length needed to obtain

TABLE I

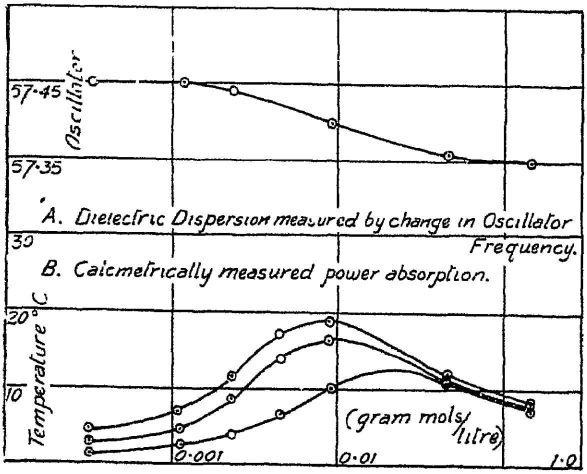
Temperature.	Frequency.
° C.	Mc./sec.
17·5	57·425
26·0	57·410
31·0	57·400
37·0	57·390

TABLE II

H_2O.	Frequency.
c c.	mc./sec.
4·7	57·52
4·8	57·50
4·9	57·48
5·0	57·46
5·1	57·44
5·5	57·38

TABLE III

$LaCl_3$ in H_2O.	Frequency.
mol.	mc./sec.
0·000327	54·45
0·00117	57·45
0·00232	57·44
0·00464	57·42
0·00926	57·40
0·0464	57·36
0·15	57·35



(Frequencies are in mc).

FIG. 2.

FIGS. 2 and 6.—Dielectric dispersion curve (A) and power absorption (B) for lanthanum chloride solutions.

λ = 500 cm.
λ = 380 cm.
λ = 212 cm.

resonant conditions both with and without the substance in position giving a determination of dielectric constant, and the width of the resonant curve, as shown by a microammeter connected to a crystal detector placed near the test condenser and symmetrical to the line, giving a determination of the power loss of the test substance. The bridges are driven by an electric motor and the detected current determined by means of a recording microammeter.

The resonator is kept in a thermostatically-controlled room. Furthermore the coupling of the resonator system to the oscillator is extremely loose, owing to the high powers available. Use is made of radiation coupling, a small dipole situated several wavelengths away from the oscillator picking up energy and transferring it to a shielded feeder, from which it is injected by re-radiation into the resonator chamber; this system effectively eliminates any "pulling of the oscillator" by the resonator.

Its range lies between 20 cm. λ and 150 cm. λ and for wavelengths greater than this a resonator must be employed making use of normal lumped inductances and capacities. The technique is standard and generally consists of placing a co-axial condenser containing the test liquid in parallel with a variable condenser across an inductance, tuning the system to resonance; a difference of condenser setting is obtained as well as a difference in the shape of the resonance curve for various test liquids placed in the co-axial condenser. The electrolytic solutions under test were contained in quartz vessels. Thus no absolute values of power loss could be obtained from direct measurement, but the significant determination of the relative width of the resonance curve could be effected.

Results.

Ionic relaxation in aqueous solutions of strong electrolytes was found, the frequency of oscillation for maximum absorption is simply related to the concentration of the electrolyte, as predicted by Debye and Falkenhagen. The dispersion of the dielectric constant has also been demonstrated. (See Table III.)

TABLE IV

Electrolyte.	Molarity (y).	Wavelength (λ). (cm.)	$\lambda y = K_\theta$.	
NaCl	0·015	1100	16·5	16·45
(Fig. 1)	0·020	850	17·0	
	0·031	530	16·4	
	0·082	200	16·4	
	0·2	80	16·0	
HCl	0·0054	530	2·86	2·85
(Fig. 3)	0·014	200	2·80	
	0·036	80	2·88	
$CaCl_2$	0·0072	1100	8·10	8·00
(Fig. 4)	0·015	530	7·95	
	0·04	200	8·00	
	0·1	80	8·00	
$SrCl_2$	0·015	530	7·95	8·00
(Fig. 5)	0·04	200	8·00	
$LaCl_3$	0·008	510	4·08	4·04
(Fig. 6)	0·0105	380	4·00	
	0·019	212	4·04	

Table IV shows the relationship for a number of monovalent, divalent and trivalent chlorides. In the last column the value of K_0 is given, from which a relaxation time may be predicted which is in general agreement with that predicted by Falkenhagen. This relationship is shown in Fig. 7.

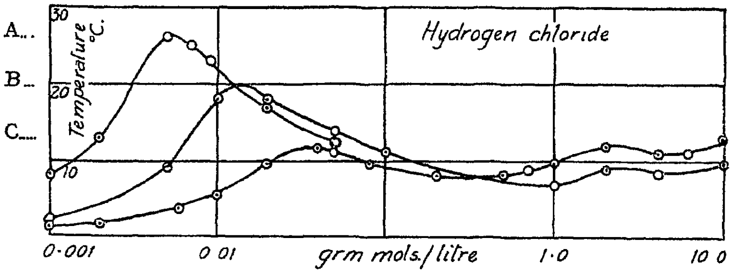
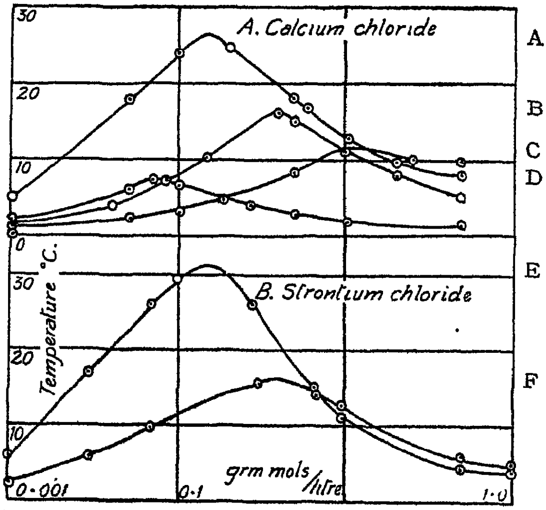


FIG. 3.—Calorimetric measurement of power absortion as a function of concentration in monovalent electrolytes.
(Wavelengths in cm.—A = 530; B = 200; C = 80.)

These results are comparable to Falkenhagen's predictions in the order of magnitude. The actual discrepancy may be due to several factors, for it should be remembered that the values obtained from Falkenhagen's equation are based on the assumptions of the Debye theory which strictly apply only to solutions more dilute than those used in the present experiments. Furthermore the temperature rise observed was in general of a high value, and although the calculations were carried out for a temperature of 25° C., some of the maxima obtained experimentally were considerably in excess of that figure. Conductivity/viscosity changes should affect the observed results, but the effect of higher temperature should be to shorten the relaxation time. It is possible that two effects come into play to explain the authors' observed higher values; for as the concentration of the solution passes through resonance the wavelength of the oscillator rises, as shown by the decrease of frequency. This can only be explained by the increase of capacity of the condenser system containing the quartz calorimeter and solution, since it is directly across the linear oscillating circuit. Since no physical movement occurs it is clear that the dielectric constant of the solution must undergo an increase, or, as was previously stated, anomalous dispersion of the dielectric constant occurs. Referring to Equation (1) it can be seen that the value D_0 may increase above the normally accepted figure for water in the region



FIGS. 4 and 5.—Calorimetric measurement of power absorption as a function in divalent electrolytes.
(Wavelengths in cm.—A=530; B=200; C=80; D=1100; E=530; F=200.

of resonance. This increase is in such a direction that the magnitude of the relaxation time will increase, and will shift the position of the maximum towards a stronger concentration. On the other hand the heating effect also increases the dielectric constant, the effects being additive. Again it is possible that the increase of relaxation time is due to the increase of the "quasi-crystalline," loose structure dimensions due to the density effect of the temperature rise.

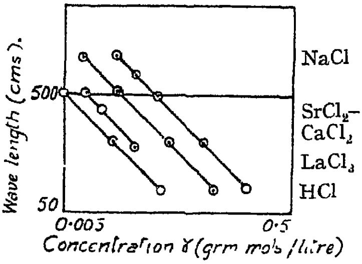


FIG. 7.—Resonance wave-length for varying concentrations of simple electrolytes, plotted logarithmically.

An examination of the curves shows that as the valency of the electrolyte increases, NaCl—$CaCl_2$—$LaCl_3$, the shape alters considerably and tends to broaden; this may be due to a restriction of motional freedom caused by the increased interionic field. A point of interest is that the rather great rises of temperature do not appear to broaden the curves of each family to any considerable extent. If the width of the curves at half-values are taken, they are found to be approximately of the same degree although the maxima may have as much as 20° C. variation between them.

TABLE V

Electrolyte.		Calculated τ value. ($\times 10^{-7}$ sec.).	Observed τ value. ($\times 10^{-7}$ sec.).
Conc. = 0·001 N.	NaCl	0·543	0·872
	HCl	0·163	0·151
	$SrCl_2$	0·505	0·847
	$CaCl_2$	0·505	0·847
	$LaCl_3$	0·484	0·645

The examination of the family of curves obtained for HCl solutions (Fig. 6) shows an effect for concentrations greater than 1 mol. which cannot be ascribed to the ionic resonance. Here it can be seen that an absorption is occurring which is not dependent on the frequencies employed, and in view of the fact that at strong concentrations the HCl molecule can exist in a non-dissociated state, it is probable that the observed effect is due to the dipole polarisation.

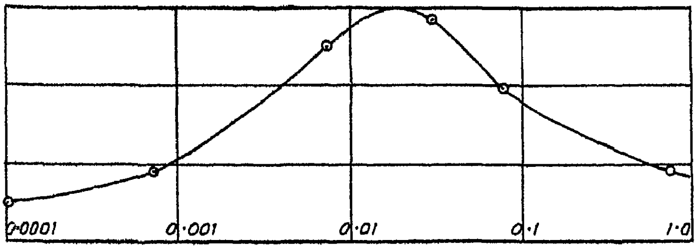


FIG. 8.—Sharpness of resonance curve as a function of concentration for sodium chloride.

Abscissæ: g. mole/l. *Ordinate:* Rel. frequency.

As a cross check, experiments were made on the linear resonator at wavelengths of 48 cm., 80 cm. and 90 cm. for solutions of $CaCl_2$ and the relation $\lambda\gamma = 8{\cdot}00$ was confirmed. Earlier experimental evidence is given on NaCl solutions where the maximum of absorption (Fig. 8) as observed in a lumped inductance—lumped capacity resonator, was obtained at a wavelength of 800 cm. for a concentration of 0·02 M., giving a K_e value of 16.[3] Glasser[4] gives a table for the maximum power absorption of aqueous NaCl solutions which agree closely to these findings. The derivation of his results appears to be somewhat obscure.

TABLE VI

Concentration for Maximum Heating, g./100 g. Solution.	Frequency, mc./sec.
0·0263	10·0
0·0526	20·0
0·0789	30·0
0·1052	40·0
0·1315	50·0

The authors are much indebted to Professor E. K. Rideal, F.R.S., for his interest and encouragement in the preliminary investigation; to Professor D. T. Harris, of the London Hospital Medical College for suggesting the problem to one of us (J. F.); and to B.X. Plastics, Ltd. for certain experimental facilities.

Department of Colloid Science,
The University, Cambridge.

---

[3] Forman, *unpublished report*, Lond. Hosp. Med. Coll., 1937.
[4] Glasser, *Medical Physics*, Chicago, 1944.

GENERAL DISCUSSION

Dr. Magat (*Paris*) said: In a recent interrogation in Paris, Professor Staudinger pointed out that marked differences exist between the melting points of paraffins and of polythenes of the same, low, molecular weight, the latter being always lower. He explains this by the assumption that a part of the ethylene molecules is added in side chains. If this assumption is correct, it would account for the existence of small losses even in unoxidised polythenes, as found by Professor Willis Jackson and his co-workers.

Professor E. Bauer (*Paris*) said: I do not think that, to explain the anomalies mentioned by Dr. Schallamach, one can speak of a Maxwell-Wagner effect, due to the conductivity of a swarm of molecules of a given sort, amongst molecules of another sort.

As a matter of fact, these swarms are not conductive. On the contrary, the fact that the variations of relaxation time, in a mixture, do not by any means follow those of the viscosity is explained by the theory which I have given here. We have here swarms of various molecules of which the lifetime depend on their size and their constitution. Similarly, the constitution of the swarms determines the crystalline fields acting on the dipoles, and therefore their apparent electric moments.

Mr. D. M. Ritson (*Oxford*) said: Dr. Schallamach's paper seems to raise considerations of extreme interest. The following might offer a simple explanation of the effect observed. In the paper of Collie, Hasted and myself, we described a very simple approach to " time of relaxation " problems. This can be generalised quite easily to the problem of a mixture of two components ((1) and (2)) in a solution.

Using the same notation and development as in our paper:

$$E' = E + \rho(P_{11} + P_{12}) + \gamma(P_{21} + P_{22})$$

$$\frac{dP_{11}}{dt} = -k_1P_{11} + k'_1 E'$$

and

$$\frac{dP_{12}}{dt} = -k_2P_{12} + k'_2E'$$

Putting $E = E_0 e^{j\omega t}$ $P_{11} = P_{011} e^{j\omega t - \psi_1}$, and $P_{12} = P_{012} e^{j\omega t - \psi_2}$

$$P_{11}(1 + j\omega T_1) = \frac{\epsilon_{s1} - \epsilon_{01}}{4\pi} + \theta_1 P_{12}$$

and

$$P_{12}(1 + j\omega T_2) = \frac{\epsilon_{s2} - \epsilon_{02}}{4\pi} + \theta_2 P_{11}$$

it can be at once seen that if θ_1 and θ_2 are equal to zero, that the result reduces to both components behaving in an " additive " manner. However, if θ_1 and θ_2 do not equal zero, " coupling " will exist between the relaxation of the two components. Dr. Schallamach's results provide strong confirmation for the existence of just such coupling between the molecules. Further, by attempting trial solutions with different values of ρ we should be able to discover which value gives the best reproduction of the experimental data.

As far as I can see, it further follows from the above that we must exercise extreme caution in attempting to find " distribution of relaxation times ". A quantitative physical meaning can only be attached to such distribution functions when we are certain that there is no co-operative polarisation, as in Debye's theory of the dielectric constant.

F. LONG CHAIN COMPOUNDS.

THE DIELECTRIC PROPERTIES OF HIGH POLYMERS.

By R. B. Richards.

Received 20*th May*, 1946.

Two reasons may be suggested for the recent interest in the dielectric properties of polymers. The first is that polymers, natural and synthetic, are widely used as insulating materials. Polythene, polystyrene, polyvinyl chloride, rubber and the phenolic thermosetting resins are familiar examples. Technical information is required on properties such as conductivity, permittivity, power factor and breakdown voltage and the manner in which these vary with frequency and temperature, on exposure to the elements or during fabrication processes such as mixing, extrusion or moulding. Published work in this field has provided a fair quantity of rather scattered and incomplete data. The second reason is that the measurement of dielectric properties provides a new tool, analogous to and complementary to the study of the absorption of higher frequency radiation in the infra-red, visible and ultra-violet range, which is already proving of value in work on the structure of polymers and the relation between chemical and physical structure and mechanical properties.

In a normal study of the relation between molecular structure and mechanical properties a deformation is applied to the outside of a matrix of molecules and the resultant effect on the individual molecules or parts of the molecules can only be assessed indirectly. By contrast, the application of an electric field to a polymer containing polar groups may be regarded as a direct application of a constraint to the molecules themselves at specific points. Attention has consequently been directed to a comparison of the relaxation times and activation energies of the processes involved in on the one hand viscous or elastic deformation of polymers, and, on the other hand, dipole orientation under alternating electric fields.

Electrical and Mechanical Relaxation Times.

It is found[1] that polymers show broader loss factor ϵ''-frequency curves than the simple Debye curve corresponding to a single relaxation time; this will be discussed in the next section. Average relaxation times τ_{DO} for dipole orientation may, however, be calculated from the position on the frequency scale of maxima in ϵ'', and from values at different temperatures of τ_{DO} figures for the free energy, heat and entropy of activation of dipole orientation may be calculated.[2] The first point to be noted is that the average relaxation times for flexible polar polymers or solutions of polar substances in flexible non-polar polymers such as polyisobutene or polythene are not vastly greater than the relaxation times for simple polar liquids or for solutions in mobile solvents such as benzene. Any attempt to relate an "internal viscosity", calculated by Debye's method from dielectric data, with the macroscopic viscosity fails completely in the case of polymeric dielectrics. Thus the ratio between the relaxation times for solutions of simple polar materials in benzene or heptane[3] and in polyisobutene[4] is of the order of 10^2 or 10^3 to 1, whereas the viscosities of the solvents are in a ratio of the order of 10^8 or more to 1. On the other hand, when a comparison is made between the activation energies of dipole orientation and the activation energies of viscous or elastic flow it is found that[5] in the case of certain flexible polymers there is a significant similarity between the values, indicating that the same molecular processes may be involved in the activation step. It may be suggested that the activation processes themselves, the twisting of quite short segments of the molecules, are the same or at least closely similar, but that viscous flow requires a higher degree of co-ordination of the segmented motion, leading to movement of whole molecules relative to one another.

The correlation between the mechanical and electrical properties of flexible polymers may be taken a stage further. The mechanical properties of polymers may be characterised by relaxation times[6] values of which may be calculated from the behaviour under high frequency oscillating mechanical deformation. In the case of certain polymers it is found that there is a close similarity between the values of the relaxation times for dipole orientation and high elastic deformation at the same temperature.[5, 7] The temperature at which a flexible polymer changes, on cooling, to a rigid or brittle state (the second order transition temperature) has been identified with the temperature at which τ_{HE} reaches a value of the order of one second, *i.e.* a value comparable with the time of duration of a typical bend-brittle or softening point test. If $\tau_{HE} \approx \tau_{DO}$ there should therefore be a relation between the temperature at which $\tau_{DO} \approx 1$ sec. and the second order transition temperature; examples of this have been pointed out.[5] A further consequence is that among polymers which are flexible at room temperature, those which have the lowest brittleness temperatures show power factor maxima, at ordinary temperatures, at the highest frequencies.

We see therefore that, in the case of flexible polymers, there is a significant amount of evidence for the identity of or at least the close similarity between the processes of dipole orientation in an A.C. field and of mechanical and, in particular, high elastic deformation. There are, however, relaxation times of dipole orientation which appear to have no parallel in mechanical phenomena. Rigid polymers, in which τ_{HE} may be regarded as being far greater than one second, may show power factor maxima at high radio frequencies; examples are unplasticised polyvinyl chloride[8] and solutions of polar materials in polystyrene.[4, 8, 9] These peaks on the power factor-frequency curves tend to be very broad. The shift with temperature, in the cases mentioned, is sluggish corresponding to low activation energies, and this suggests that some process other than the twisting of the main polymer chain takes place in the activation process. Movement of side groups relative to the main chain is a possible cause of such subsidiary maxima. Much work clearly remains to be done on the mechanism of dielectric loss in rigid amorphous polymers.

---

[1] Fuoss and Kirkwood, *J.A.C.S.*, 1941, 63, 385.

[2] Kauzmann, *Rev. Mod. Physics*, 1936, 7, 34; Frank, *Trans. Faraday Soc.*, 1936, 32, 1634; Powell and Eyring, *Advances in Colloid Science*, Interscience Publishers, New York, 1942.

[3] Whiffen and Thompson, *Trans. Faraday Soc.*, 1946, 42A, 114.

[4] Plessner and Richards, *Trans. Faraday Soc.*, 1946, 42A, 206.

[5] Tuckett, *Trans. Faraday Soc.*, 1944, 40, 448.

[6] *E.g.*, Alexandrov and Lazurkin, *Acta Physicochimica U.R.S.S.*, 1940, 12, 647.

[7] Ponomarev, *J. Tech. Phys. U.R.S.S.*, 1940, 10, 587.

Distribution of Relaxation Times in Polymers.

Power factor-frequency curves which are broader than the simple Debye curve for a single relaxation time are given by all polymeric dielectrics, even by a simple system such as a solution of a small molecule ester in a flexible polymer such as polyisobutene.[1] It has indeed been pointed out[1] that a relaxation time distribution is an inherent property of a dielectric composed of long molecules. That the effect of length of chain on relaxation time distribution shows itself at an early stage is demonstrated by the decrease of the α factor * from unity for solutions in heptane[3] (M.W. 100) to 0·68 for a solution in medicinal paraffin[10] (M.W. of the order of 300-400) the value for a solution in polyisobutene[4] (M.W. ca. 100,000) being 0·68 also.

The spread of relaxation times may be further increased by any inhomogeneity of molecular architecture or of physical state. Thus an increased breadth of distribution may be expected if there are more than one type of dipole, as in a copolymer or, for example, in chlorinated polythenes,[11] or if the dipoles are distributed at irregular intervals along the chain. The broad spread of relaxation times in chlorinated polythene containing 40 % chlorine compared with the spread in neoprene[12] with about the same chlorine content but with a more regular structure, may be given as an example. Similarly we may expect a broader spread in a semi-crystalline polymer than in a completely amorphous polymer; the comparison[4] of the properties of solutions of esters in polythene and in polyisobutene is relevant here.

One final point regarding the distribution of relaxation times may be mentioned. As the properties of a flexible polymer are changed so that it approaches the rigid state, as for example by a reduction of temperature or of plasticizer content,[8] or by chemical modification such as the introduction of chlorine into the polythene molecule,[11] the value of α is found to decrease, a value of zero (corresponding to a flat loss factor-frequency curve or an infinite spread of relaxation times) being approached at the temperature, plasticizer content, etc. corresponding to the transition to a rigid state.

Filled Polymers.

Polymers may be used as dielectrics in combination with second components such as carbon black or mineral fillers or pigments. In an interpretation of the dielectric properties of such a heterogeneous dielectric Wagner's[13] extensions of Maxwell's equations must be used. According to these equations the properties of a dielectric compound of a dispersion of discrete particles of one material in a second material are related to the volume fractions, permittivities and conductivities of the components. This treatment has led to an interesting use [12, 14] of dielectric measurements in determining whether a filler behaves as an inert diluent or interacts with the polymer molecules. If the filler is entirely inert the relaxation times of dipole orientation of the polymer molecules would be the same in the pure polymer and in the filled material. If, however, there is interaction as with a reinforcing filler, the relaxation times in the pure polymer and in the compound would differ.

---

[8] Fuoss, *J.A.C.S.*, 1941, **63**, 369.

[9] Frank and Jackson, *Trans. Faraday Soc.*, 1940, **35**, 440.

[10] Powles and Jackson, *Trans. Faraday Soc.*, 1946, **42A**, 101.

[11] Oakes and Richards, *Trans. Faraday Soc.*, 1946, **42A**, 197.

[12] Schneider, Carter, Magat and Smyth, *J.A.C.S.*, 1945, **67**, 959.

[13] Wagner, *Archiv. Elektrotechnik*, 1914, **2**, 371.

* Fross and Kirkwood[1] have shown that experimental loss factor-frequency curves for polymers may be represented by an empirical equation including a parameter α which has a value of unity for a single relaxation time, but approaches zero as the spread of relaxation times increases.

Summary.

A brief review is made of some features of the dielectric properties of high polymers with particular reference to points raised in the papers of Carter, Magat, Schneider and Smyth,[14] of Oakes and Richards [11] and of Plessner and Richards [4] presented at this General Discussion.

Measurement of dielectric properties provides a new tool for the study of polymer structure and properties. In the case of flexible polymers some success has been achieved in the identification of the activation processes and of relaxation times of dipole orientation and of high elastic deformation, but in the case of rigid polymers the position is more complex. Polymeric dielectrics show broad relaxation time distributions; this is an inherent property of their structure, but may be accentuated by chemical or physical inhomogeneity as in a copolymer or in a semi-crystalline polymer. The distribution in a flexible polymer broadens as the conditions (temperature, plasticiser content, etc.) are changed towards the point at which flexibility is lost in a transition of the second order. The dielectric properties of an inhomogeneous dielectric such as a filled polymer may be analysed in terms of the Wagner-Maxwell equations; here dielectric measurements give an indication of any interaction between polymer and filler as in the case of a reinforcing carbon black in rubber.

Imperial Chemical Industries Research Dept.,
Alkali Division,
Northwich.

---

[14] Carter, Magat, Schneider and Smyth, *Trans. Faraday Soc.*, 1946, **42A**, 213.

THE DIELECTRIC PROPERTIES OF CHLORINATED POLYTHENES AT RADIO FREQUENCIES.

BY W. G. OAKES AND R. B. RICHARDS.

Received 3rd March, 1946.

Attention has recently been directed to the interrelation of the mechanical and the electrical properties of polymeric materials, and it has been suggested that the orientation of sections of the chain molecules which occur on the one hand during high elastic deformation and on the other hand under the influence of an applied electric stress are similar and can be characterised by similar relaxation times or activation energies. Changes in the chemical structure of a polymer, such as the introduction of polar groups or of bulky side chains, lead to changes in the properties of the individual molecule, such as the readiness with which it can coil up and uncoil by rotation about the bonds connecting the atoms of the main chain, and also to changes in the intermolecular forces. An increase in the concentration of a polar side group can in general be expected to increase intermolecular forces and decrease molecular flexibility, and hence to increase both mechanical and electrical relaxation times and their associated activation energies.

The thermoplastic polymers made by the progressive substitution chlorination of polythene[1] form a suitable series for a study of this point. Starting with a material composed of molecules with virtually no dipoles other than the weak C—H group, a continuous series of polymers is available with increasing chlorine content up to over 70 % by weight (corresponding to approximately one chlorine atom per carbon atom), but with the same carbon skeleton. This increase in chlorine content leads to changes in the properties at normal temperature from the flexible but fairly stiff state characteristic of polythene first to a rubbery material and then to a rigid brittle plastic. As will be shown in this paper, the dielectric properties at radio frequencies change in a manner which indicates a similar progressive increase in the relaxation time.

In addition, by altering the conditions of chlorination, polymers may be made which, although they have the same chlorine content and, being based on the same polythene, have the same chain length, have nevertheless different mechanical properties. This difference, due in part at least to an alteration of the distribution of the chlorine atoms along the carbon-carbon chain, is associated with differences in electrical properties which can be attributed to the same cause.

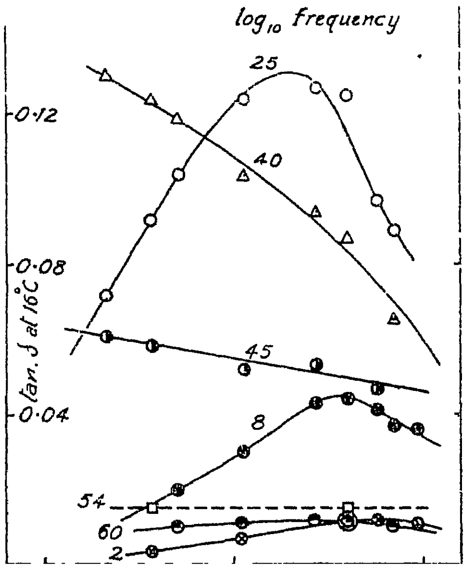


FIG. 1.—Power factor—frequency curves at 16° C. for hot process chlorinated polythenes. (⊕ 2, ● 8, ○ 25, △ 40, ◑ 45, ☐ 54, ⊖ 60 % Cl.)

Experimental.

Preparation of Chlorinated Polythene Samples.—The samples used in this work were all prepared by Mr. L. Furnival. Two series were made from the same polythene ("Alkathene" 7), the average chain length of which was about 1300 carbon atoms. The first series was made by chlorination of a solution of polythene in carbon tetrachloride at 65-70° C. This series will be referred to as the hot process series. The second series was made by chlorinating a suspension of finely divided polythene in carbon tetrachloride at 30-40° C. This series will be referred to as the cold process series. There is no evidence of any degradation of the basic carbon-carbon skeleton during chlorination.

Measurement of Dielectric Properties.—Hartshorn-Ward N.P.L. test sets[2] (H. Tinsley & Co.) were used over the frequency range 10^5-10^8 cps. One set, at Manchester University, had been modified for variable temperature work by Mr. Plessner and Mr. Reddish and will be described in a forthcoming paper. The test discs were cut from 1 mm. sheet which was moulded at 100-130° C., the moulding temperature being chosen to suit the softening point of the sample.

[1] Fawcett and I.C.I. Ltd., *B.P.* 481,515 (1936).

[2] Hartshorn and Ward, *J.I.E.E.*, 1936, **79**, 597.

Experimental Results.

Effect of Chlorine Content on the Power Factor-Frequency Curve.—Fig. 1 shows plots of tan δ against the frequency *f*, at room temperature (*ca.* 16° C.) for members of the hot process series. As chlorine is introduced, the first effect is a marked rise in the power factor over the whole range from 10^5-10^8 cps., with a pronounced maximum which moves to lower frequencies as the chlorine content increases. Further introduction of chlorine leads to a fall in tan δ at these frequencies, an effect which would be expected as a result of the movement of the tan δ peak to frequencies below the range studied in this work. Fig. 2 shows the corresponding permittivity-frequency curves, for some of the samples; these show the usual fall in ε′ over the frequency range in which tan δ shows a peak.

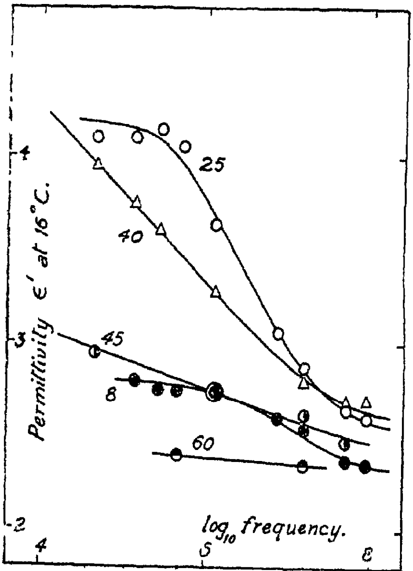


FIG. 2.—Permittivity-frequency curves at 16° C. for hot process chlorinated polythenes. (Symbols as in Fig. 1.)

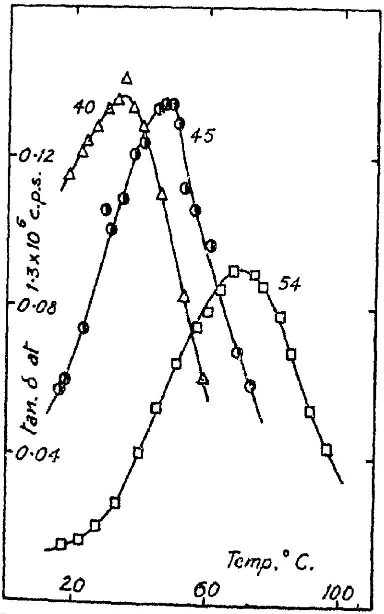


FIG. 3.—Power factor—temperature curves for hot process chlorinated polythenes at $1{\cdot}3 \times 10^5$ cps. (Symbols as in Fig. 1.)

Variable Temperature Measurements.—The peak in tan δ at room temperature appears to lie below the frequency range used in this work in the case of the samples of the hot chlorinated series of chlorine contents of 40 % and above. Some measurements at higher temperatures were therefore made; the results for a frequency of $1{\cdot}3 \times 10^6$ are shown in Fig. 3. As the chlorine content is increased the temperature at which tan δ is a maximum rises, but the height of the peak decreases.

Effect of Chlorination Conditions.—Fig. 4 shows the values of tan δ at $1{\cdot}6 \times 10^7$ cps. and at room temperature as a function of the chlorine content for the hot and the cold chlorinated series. For a given chlorine content up to about 40 % the hot chlorinated series have values of tan δ more than twice those of the cold chlorinated series. This could be caused merely by a shift of the peak in tan δ to lower frequencies in the case of the cold chlorinated series, leaving the value of tan δ_{max} or the area under the tan δ-frequency curve the same. That this is not the explanation is shown by a comparison of the tan δ-frequency curves for samples of the two series with the same chlorine content. It is found, in the case of the samples with 8 % chlorine for example, that tan δ is greater in the case of the hot chlorinated sample over the whole frequency range 10^3-10^8 cps. The frequencies at which tan δ has a maximum value differ little, being about $1 \cdot 5 \times 10^7$ cps. for the hot and $0 \cdot 8 \times 10^7$ cps. for the cold chlorinated sample.

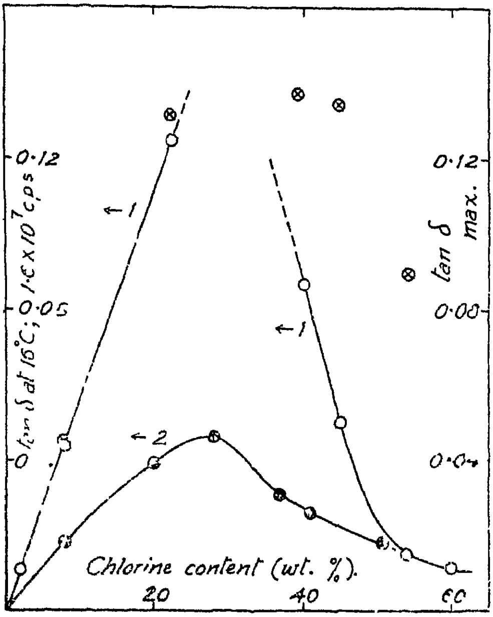


FIG. 4.—(*a*) Power factor (tan δ at 16° C. and $1 \cdot 6 \times 10^6$ cps.) of:

(1) Hot process chlorinated polythenes ○
(2) Cold process chlorinated polythenes ●

as a function of chlorine content. (tan δ for original polythene = 0·00025.)

(*b*) Peak power factor (tan $\delta_{max.}$) for hot process samples ⊗. (0-25 % Cl from variable frequency measurements at room temperature, 40-54 % from variable temperature measurements at $1 \cdot 3 \times 10^6$ cps.)

TABLE I.—POWER FACTORS OF HOT AND OF COLD PROCESS CHLORINATED POLYTHENES WITH 8 % CHLORINE AT CA. 16° C.

Frequency (cps.).	Power Factor (tan δ).	
	Hot.	Cold.
$2 \cdot 6 \times 10^5$	0·020	0·012
$1 \cdot 35 \times 10^6$	0·030	0·015
$7 \cdot 5 \times 10^6$	0·043	0·022
$1 \cdot 6 \times 10^7$	0·044	0·021
$3 \cdot 3 \times 10^7$	0·041	0·018
$9 \cdot 0 \times 10^7$	0·036	0·016

Discussion of Results.

Electrical and Mechanical Properties.—The power factor-frequency curves (Fig. 1) for the hot chlorinated series indicate that as the chlorine content is increased there is, at first at least, an increase in the peak in tan δ and in the area under the tan δ-frequency curve, which may be regarded [8] as a measure of the product of total number of dipoles and the square of the average dipole moment; this is discussed in detail later. In addition there is a shift of the peak in tan δ to lower frequencies. This explains the results of Fig. 4, the initial rise in tan δ at a high frequency, followed by a fall. This shift in the peak to lower frequencies is an indication of an increase in the average relaxation time τ_{D0} of dipole orientation from the very low value (for a polymer) of about 2×10^{-9} seconds which is characteristic of the unsubstituted polythene chain to a value of about $1 \cdot 5 \times 10^{-7}$ secs. at 25 % chlorine. More chlorine leads to even higher values at room temperature, and the temperature required to give a peak in tan δ at a given frequency, *i.e.* at which the average relaxation time has a given value, rises as the chlorine content rises.

---

[8] Sillars, *Proc. Roy. Soc.*, *A*, 1939, 159, 66.

This increase in τ_{D0} with chlorine content may be considered in connection with the parallel effect of an increase in the concentration of chlorine atoms along the polymer chain on the mechanical properties. This is summarised for the hot process samples in Table II.

TABLE II.

INFLUENCE OF CHLORINE CONTENT ON MECHANICAL PROPERTIES OF HOT PROCESS CHLORINATED POLYTHENES.

Chlorine Content (Wt. %).	Ratio of No. of Cl Atoms to C Atoms.	Structure at Room Temp.	Physical Form at Room Temp.	Brittleness (a) Temp. °C.	Softening (b) Pt. °C.
2	1 in 120	Mainly crystalline	Flexible, similar to polythene	—	82
8	1 in 28	Crystalline	More flexible	Below −70	69
25	1 in 7·8	Mainly amorphous	Rubbery	,,	Below 20
40	1 in 3·9	Amorphous	Soft and flexible, "lazy"	,,	*Ca.* 20
45	1 in 3·2	,,	Flexible, leathery	*Ca.* −20	30
54	1 in 2·2	,,	Rigid	*Ca.* +20	52
60	1 in 1·8	,,	Rigid, brittle	*Ca.* +40	67

(a) This is the temperature at which 50 % of a number of strips of 1·6 mm. sheet break when bent rapidly round a mandrel of radius 8·0 mm. The figure for a high molecular weight polythene is below −70° C.

(b) This, the modified Vicat softening point used for polythene, is the temperature at which a square ended needle, of area 1 sq. mm., with a load of 1000 gm., achieves a penetration of 1 mm. into a 3 mm. disc of the specimen, the temperature being raised at a rate of 50° C. per hour from a temperature more than 20° C., below the softening point. The figure for "Alkathene" 7, used as the basis of the chlorinated derivatives used in this work, is about 90° C.

The first effect of chlorination is to reduce the ability of the polymer chains to crystallise in the paraffin-like structure found in polythene.[4] This leads to a reduction in the softening point and the crystal melting-point, the chlorine atoms still being so widely spaced that the flexibility of the polymer chain is not markedly reduced so that the products are flexible at low temperatures. The crystalline paraffin X-ray pattern disappears in the hot chlorinated series at a chlorine content of around 35 %.[5] When the chlorine content is increased above 40 % the products, although now amorphous, become increasingly stiff and both brittleness temperatures and softening points rise. This may be regarded as an increase in the second order transition temperature or high elasticity temperature T_{HE} as the chlorine content increases; the change from a rigid solid to a rubber is not sharp, however, and it is difficult to allot precise values to T_{HE}. The second order transition temperature has been identified[6] with the temperature at which the average relaxation time T_{HE} which characterises high elastic deformation, reaches a value of about 1 second. This second order transition temperature is in the neighbourhood of room temperature for a sample of about 50 % chlorine content, *i.e.* T_{Bm} is *ca.* 1 second at 16° C. for this material. The increase in the corresponding electrical relaxation time τ_{D0} with increase in chlorine content is not inconsistent with the suggestion that a value of *ca.* 1 second at room temperature would also be reached at a chlorine content of *ca.* 50 % (Fig. 5). The approximately parallel increase in the softening and brittleness temperatures of the hot process samples with the higher chlorine contents and the temperature at which tan δ has a maximum at a frequency of $1{\cdot}3 \times 10^5$ cps., *i.e.* at which $\tau_{D0} \approx 10^{-6}$ secs., may also be noted (Fig. 3 and Table II).

---

[4] Bunn, *Trans. Faraday Soc.*, 1939, **35**, 482; 1945, **41**, 317.

[5] Bunn, *unpublished.*

[6] Tuckett, *Trans. Faraday Soc.*, 1942, **38**, 310.

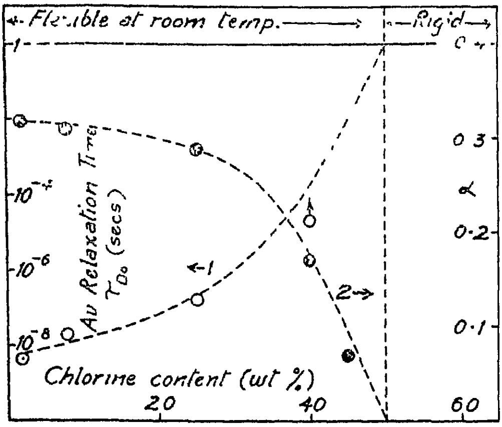


FIG. 5.—Influence of chlorine content on: (1) Average electrical relaxation time τ_{D0} at room temperature (16° C.). (2) Parameter α measuring breadth of relaxation time distribution at 16° C. (α = 1 for a single relaxation time.)

Relaxation Time Distributions and Average Dipole Moment.—The curves of Fig. 1 are all broader than the simple Debye curves with a single relaxation time. As a quantitative measure of the breadth of relaxation time distribution, values of α in the empirical equation of Fuoss and Kirkwood [7]

$$\text{Loss Factor } \epsilon'' = \epsilon' \tan \delta = A \operatorname{sech} (\alpha \log f_{\max}/f) \quad (1)$$

have been calculated. This equation fits fairly well the observed loss factor-frequency curves for the hot process samples containing 2 %, 8 % and 25 % chlorine. The values of α for these samples, together with more approximate values for the samples with 40 % and 45 % chlorine based on the limited range of values of tan δ shown in Fig. 1, and on a value of $\epsilon''_{\max}$ estimated from the variable temperature results, are plotted as a function of chlorine content in Fig. 5. The value of α, which is unity for a single relaxation time and zero for an infinitely wide distribution, is seen to decrease at room temperature (16° C.) as the chlorine content increases. Fuoss and Kirkwood observed that in a series of plasticised polyvinyl chloride samples, the value of α decreased, approaching zero, as the plasticiser content decreased at constant temperature, or as the temperature decreased at constant plasticiser content, that is, as the sample became more and more rigid or as the internal melting-point or second order transition was approached. This effect seems to be shown in the chlorinated polythenes, the increase in chlorine content being analogous to a decrease in plasticiser content.

It may be noted that the values of α for the chlorinated polythenes are similar to the values found for plasticised polyvinyl chloride samples by Fuoss and Kirkwood,[7] but the value for the hot process sample with 40 % of chlorine is less than the value (0·45 at 20° C.) for neoprene [8] which has an almost identical chlorine content but a more regular chemical structure. The peak in tan δ for neoprene at 20° C. lies, in addition, at a much higher frequency (*ca.* 5×10^7 cps.) than the peak (below 10^4 cps.) for the corresponding chlorinated polythene.

From the area under the loss factor-frequency curve, or from the values of the peak loss factor and of α, and from the known number of chlorine atoms per molecule, it is possible to calculate the average effective dipole moment $\overline{\mu}$ of a C—Cl link in chlorinated polythene. Using the equations[3]

---

[7] Fuoss and Kirkwood, *J. Am. Chem. Soc.*, 1941, 63, 385.

[8] Schneider, Carter, Magat and Smyth, *J. Am. Chem. Soc.*, 1945, 67, 959.

$$\text{Area} = \int_0^\infty \epsilon' \tan \delta \,.\, d \log_e f = \frac{2\pi^2}{27} \frac{(\epsilon_0 + 2)^2}{kT} \overline{\mu}^2 n \quad . \quad . \quad (2)$$

or

$$\pi \epsilon''_{\max} = \frac{2\pi^2}{27} \frac{(\epsilon_0 + 2)^2}{kT} \overline{\mu}^2 n \quad . \quad . \quad . \quad . \quad (3)$$

where n is the number of chlorine atoms per c.c. of dielectric, values of $\overline{\mu}$ have been calculated for the hot process samples with 2 %, 8 % and 25 % chlorine and the cold process sample with 8 % chlorine (Table III).

TABLE III.

Average Dipole Moments per C—Cl Link, Calculated from Power Factor-Frequency Curves.

Sample.	Average Dipole Moment (Debye Units).
2 % hot process .	1·56
8 % hot process .	1·44
25 % hot process .	1·38
8 % cold process .	1·02
Neoprene[8] . .	1·99[(a)], 2·15[(b)]
Plasticised polyvinyl chloride . .	2·1-2·8[(a)]

[(a)] Calculated by method of Fuoss and Kirkwood.[7]

[(b)] Calculated from equation (3).

These values may be compared with some values for simple chlorinated paraffins given in the 1934 Faraday Society list[9] (Table IV).

A comparison of Tables III and IV suggests that possible explanations of the low values of the average dipole moment per C—Cl link in the chlorinated polythenes, compared with, for example, neoprene and plasticised polyvinyl chloride, is the presence of some of the chlorine in —CCl_2— groups or an adjacent carbon atom as —CHCl—CHCl— or —CHCl—$CHCl_2$ groups. The effect of the distribution of chlorine atoms on the electrical and mechanical properties is discussed in the next section. The decrease in tan $\delta_{\max}$ as the chlorine content increases in the variable temperature measurements on the high chlorine content samples (Fig. 3) may also be attributed in part to a fall in the effective average dipole moment per chlorine atom due to the increased proportion of groups such as —CCl_2— or —CHCl—CHCl—.

TABLE IV.

Dipole Moments of Simple Chlorinated Paraffins.

Type.	Material.	Dipole Moment.	Dipole Moment per C—Cl Link.
Single CHCl group . .	CH_3Cl	1·56-2·0	1·56-2·0
	C_2H_5Cl	1·90-2·05	1·98-2·05
	$CH_3 . CHCl . CH_3$	2·04-2·18	2·04-2·18
	$C_3H_7 . CHCl . C_3H_7$	2·0 -2·04	2·0 -2·04
Single CCl_3 group . .	CH_2Cl_2	1·48-1·62	0·74-0·81
	$CH_3 . CHCl_2$	1·8 -2·05	0·90-1·02
	$C_2H_5 . CHCl_2$	2·06	1·03
	$CH_3 . CCl_2 . CH_3$	2·0 -2·18	1·0 -1·09
Adjacent CHCl groups .	$CH_3 . CHCl . CH_2Cl$	1·85	0·92
Adjacent CHCl and CCl_2 groups . . .	$CH_2Cl . CHCl_2$	1·15-1·55	0·57-0·77

Distribution of Chlorine Atoms along the Chain.—In the hot chlorination the polythene molecules, being in solution in a good solvent, are likely to be in an extended form. Chlorination, in the early stages, will probably occur randomly along the chain, all hydrogen atoms being equally likely to be replaced by chlorine atoms. If this were to persist throughout the chlorination, it would be possible to calculate statistically the proportions of carbon atoms in the forms —CH_2—, —CHCl— and —CCl_2— at any stage of chlorination. Steric and other directive effects will, however, become important after the earliest stages, and any purely statistical calculation of the chlorine distribution would be misleading. Some indication of the distribution of the chlorine atoms along the chain is given by the infra-red data of Thompson[10] on a series of hot process samples, which shows that as the chlorine content increases the intensity of a band at 660 $cm.^{-1}$ present in polyvinylidine chloride and probably due to the —CCl_2— group, increases relative to that of a band at 620 $cm.^{-1}$, present in polyvinyl chloride and probably due to the —CHCl— group, and to that at 1460 $cm.^{-1}$, due to —CH_2—. In addition there is a decrease in the intensity of a band at 722 $cm.^{-1}$ found in paraffins containing, and probably characteristic of, four or more adjacent methylene groups (—CH_2—CH_2—CH_2—CH_2—). This band finally disappears at a chlorine content of about 50 % in the hot chlorinated series. As noted above, the X-ray pattern typical of polythene crystallites which is also an indication of lengths of unchlorinated chain also becomes weaker as chlorination proceeds.

---

[9] *Trans. Faraday Soc.*, 1934, **30**, Appendix.

In the cold chlorinated series random chlorination is not to be expected. The particles of polythene, although swollen by the carbon tetrachloride, will be partly crystalline[11] and the attack of chlorine will probably be concentrated on the amorphous regions. As chlorine is added, the lengths of unchlorinated molecule and the proportion of crystalline material decreases but the distribution of polar groups in the products of intermediate chlorine content will be expected to be different from that in the hot chlorinated series. This difference may be represented diagrammatically thus :

—CHCl—CHCl—CH_2—CHCl—CH_2—CH_2—CCl_2—CH_2—CHCl—CH_2—CHCl—CH_2—

Hot Chlorination

—CH_2—CH_2—CH_2—CH_2—CH_2—CH_2—CCl_2—CHCl—CHCl—CCl_2—CH_2—CHCl—

Crystallite of
original polythene.

Amorphous region of
original polythene

Cold Chlorination.

In addition there is a difference in the chlorine contents of the component molecules of cold chlorinated samples, the molecules near the surface of the suspended particles being more heavily chlorinated than those in the middle. The net result will be that at any chlorine content (except possibly the very highest) a sample made by the cold process will contain greater lengths of unchlorinated chain. The paraffin-like X-ray pattern is present at normal temperatures up to a chlorine content of about 55 %, *i.e.* an average of one chlorine atom on every other carbon atom, in the cold process series, compared with the value of 35 % for the hot process series.[5] The greater proportion of unchlorinated chain means that there must be a greater concentration of chlorine atoms, and hence a greater proportion of CCl_2 groups or adjacent chlorinated carbon atoms leading, by analogy with the figures in Table IV, to a lower effective average moment per C—Cl group and to the lower power factors of the cold chlorinated series (Fig. 4).

Such a difference in chlorine atom distribution would be expected to affect the mechanical properties of the polymer. The higher proportion of lengths of unchlorinated chain, leading to higher crystallinity, will be expected to increase the hardness and softening point of the products, and the high chlorine concentration on the remainder of the chains may be expected to reduce the flexibility, by steric hindrance or inter-chain action, of the chlorinated section. Thus we find that, for a given chlorine content, the cold process samples have the higher softening point. Comparative figures, measured by the needle penetration test described under Table II, are shown in Table V.

---

[10] Thompson, *Trans. Faraday Soc.*, 1945, 41, 246.
[11] Richards, *ibid.*, 1946, 42, 10.

TABLE V.

Softening Points of Hot and Cold Process Samples.

Chlorine Content (Wt. %).	Softening Point °C.	
	Hot Process.	Cold Process.
0 (Polythene)	90	
8	69	79
25	Below 20	—
28	—	65
40	Below 20	69
50	40	77

Summary.

The power factors and permittivities of chlorinated polythenes containing up to 60 % of chlorine have been measured at room temperture over the frequency range 10^4-10^8 cps. Two series, made by the chlorination of polythene in solution (hot process) and in suspension (cold process), were examined. As the chlorine content increased up to 40 % the height of the peak in tan δ rose and the peak moved to lower frequencies, being at *ca.* 5×10^7 cps. for the lowest chlorine content and below 10^4 cps. for a product containing 40 % of chlorine in the hot process series. In variable temperature experiments at a fixed frequency on samples of 40 % chlorine and above, the temperature at which tan δ reached a maximum rose as the chlorine content increased but the height of the peak decreased. The hot process samples had higher power factors at a given frequency than the cold process samples.

The increase of the average dipole orientation relaxation time as chlorine content increases, indicated by the tan δ-frequency curves, is in accord with the change from a flexible or rubbery to a rigid state. The breadth of the relaxation time distribution increases at room temperature as the chlorine content rises. Average dipole moments per C—Cl link have been calculated from the tan δ-frequency curves; the values are lower than the values for ethyl chloride, neoprene or plasticised polyvinyl chloride and this is interpreted as being due to the presence of —CCl_2— or —CHCl—CHCl— groups, the effective average moment being reduced by dipole interaction. The difference in electrical properties between the hot process and the cold process samples, and the higher softening points of the latter, are discussed in terms of the distribution of chlorine atoms along or among the polymer molecules.

The authors wish to thank members of this Research Department for assistance with the work described in this paper. Mr. D. Whittaker gave much advice on the selection of samples and supplied information on the properties of chlorinated polythenes, and Miss M. Brazendale and Mr. C. Peover assisted with the experimental work.

Imperial Chemical Industries, Limited,
Research Department,
Alkali Division,
Northwich.

DIPOLE ORIENTATION IN SOLUTIONS OF ESTERS IN POLYISOBUTENE AND POLYTHENE.

By K. W. Plessner and R. B. Richards.

Received 1st March, 1946.

Pure hydrocarbons have very low power factors over a wide range of frequencies. When polar groups are introduced, either as a chemical modification of the hydrocarbon molecule or in the form of polar molecules in solution, the dielectric has an appreciable power factor, the value of which and the frequency at which it is a maximum being governed by the concentration and dipole moment of the polar group and the relaxation times which characterise the orientation of the dipoles in an electric field. These relaxation times are in turn characterised by entropies and energies of activation of dipole orientation, and will depend both on the size, shape and polarity of the polar group and on the properties of the environment in which orientation is occurring. In the case where the dipole is part of a dissolved molecule, therefore, we expect different average relaxation times and distributions of relaxation times for solutions of the same polar solute in different hydrocarbons.

Debye's original treatment of power loss applies to solutions of small, approximately spherical, polar molecules in a liquid solvent, and experimental data on dielectrics of this type has been published; [1] the emphasis here has naturally been on the correlation of the values of the viscosity of the solvent and the radius of the solute molecules with the values of the same quantities calculated from Debye's theory. The work was extended to solids by Jackson,[2] Sillars [3] and Pelmore and Simons,[4] who measured over a range of frequencies and temperatures the power factor of solutions of linear esters in solid paraffin wax, and Fröhlich [5] has shown how these results agree with his theoretical calculations of the relaxation times for such a system. Another solid hydrocarbon solvent which has been examined is polystyrene; [6, 7] here it is found that although the polystyrene is itself a rigid solid at normal temperatures, a dissolved polar solute leads to a power loss with a maximum at radio frequencies.

Polyisobutene and polythene are representative of two further types of dielectric. Polyisobutene is, at normal temperatures, an amorphous rubbery solid, while polythene is partly crystalline and is intermediate in rigidity between polyisobutene and polystyrene. The differences in structure and mechanical properties suggest that there will be characteristic differences in the dielectric properties of solutions of polar materials in these hydrocarbons; data on this subject is presented in this paper.

Experimental Results.

(*a*) **Measurement of Power Factor.**—An N.P.L. Dielectric test set (H. Tinsley & Co.), of the type described by Hartshorn and Ward,[8] was used over the frequency range 10^4-10^8 cps. for the bulk of the measurements.

---

[1] *E.g.* Mueller, *Naturwiss.*, 1938, **18**, 202.

[2] Jackson, *Proc. Roy. Soc.*, *A*, 1935, **150**, 197.

[3] Sillars, *ibid.*, 1938, **169**, 66.

[4] Pelmore and Simons, *ibid.*, 1940, **175**, 253.

[5] Fröhlich, *Proc. Physic. Soc.*, 1942, **54**, 422; *J.I.E.E.*, 1944, **94**, Pt. I, 456.

[6] Frank and Jackson, *Trans. Faraday Soc.*, 1940, **35**, 440.

[7] Fuoss, *J.A.C.S.*, 1941, **63**, 369.

[8] Hartshorn and Ward, *J.I.E.E.*, 1936, **79**, 597.

The set was adapted for use at a range of temperatures from −80° C. to +80° C., over the frequency range 10^4-10^7, by the addition of an extra condenser in parallel with the main condenser of the measuring circuit. This extra condenser was installed in a Dewar vessel through which dry air, cooled or heated to the appropriate temperature, was passed.

Some figures at $2{\cdot}8 \times 10^9$ cps. were measured by the resonant cavity method.[9]

(*b*) **The Power Factor of a Polyisobutene-Dioctyl Phthalate Mixture.**—Dioctyl phthalate was chosen as the solute for the comparison of the loss in polyisobutene and polythene because its low vapour pressure made possible the preparation of homogeneous mixtures of known concentration on rubber mills, and because the resultant dielectrics showed their maximum power factor at a frequency convenient for use with the N.P.L. test set. Fig. 1 shows tan δ-frequency curves for a solution of 2 % by weight of dioctyl phthalate in polyisobutene (M.W. *ca.* 100,000) at a series of temperatures. The curves shown here and in Figs. 2 and 3 represent the power factors due to the solute molecules, the power factor of the solvent having been subtracted from the observed values.

The curves obtained at these temperatures are very similar in shape, with a slight decrease in the peak power factor tan δ_{max} as the temperature increases. The curves are, however, broader than a normal Debye curve with a single relaxation time; this point will be discussed in more detail later.

(*c*) **The Power Factor of a Polythene-Dioctyl Phthalate Mixture.**—Fig. 2 shows a similar series of curves for a 2 % solution of dioctyl phthalate in polythene ("Alkathene" 20). These differ from the corresponding polyisobutene curves in two respects: (1) at any temperature the power factor maximum occurs at a higher frequency, and (2) the curves at the lower temperatures are broader, with lower values of tan δ_{max} than those for the higher temperature, or than those for polyisobutene-dioctyl phthalate mixtures.

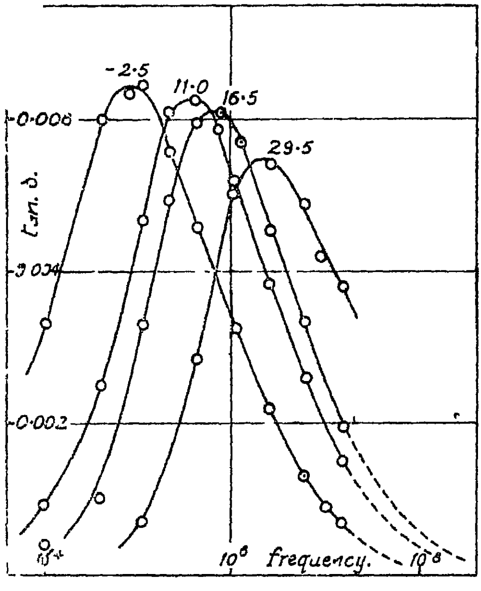


FIG. 1.—Power factor-frequency curves for 2 % dioctyl phthalate in polyisobutene.

(*d*) **The Power Factor of a Polystyrene-Dioctyl Phthalate Mixture.**—For comparison with the results for the flexible polymers polyisobutene and polythene, some power factors of polystyrene ("Transpex" II) containing dioctyl phthalate, have been measured. Power factor-frequency curves for 2 % of dioctyl phthalate in polystyrene showed maxima, but the value of tan δ_{max} was much lower than that for the flexible polymers, and a higher ester content was therefore examined. Fig. 3 shows values of the power factor for a 10 % solution of dioctyl phthalate in polystyrene at a series of temperatures. The value of tan δ_{max} increases as the temperature rises and the frequency f_{max} at which tan δ is a maximum also increases, but the increase of f_{max} with temperature is much more sluggish than in the cases of polyisobutene or polythene. The curves are markedly unsymmetrical.

---

[9] Horner, Taylor, Dunsmuir, Lamb and Jackson, *J.I.E.E.*, 1946, 93, Pt. III, 53.

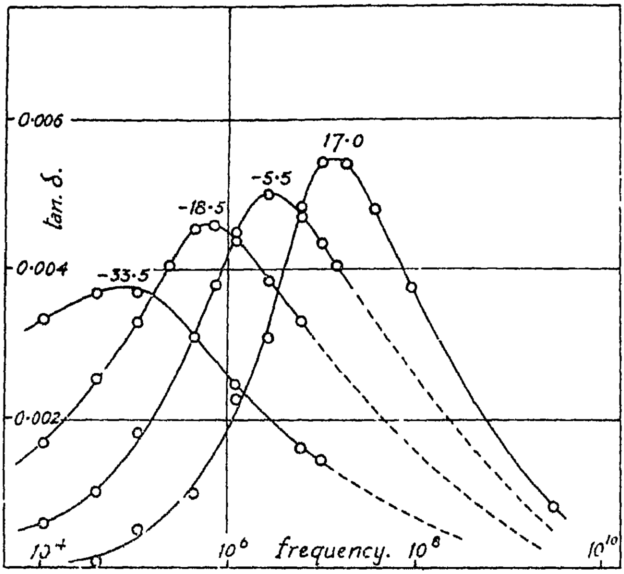


FIG. 2.—Power factor-frequency curve for 2 % dioctyl phthalate in polythene.

(e) The Power Factor of Solutions of Other Esters in Polythene.—The relaxation time of orientation of the polar molecules will be a function not only of the properties of the solvent hydrocarbon, but also of the size and nature of the solute molecules. Tan δ-frequency curves have been determined at room temperature (*ca.* 16° C.) for a series of solutions of

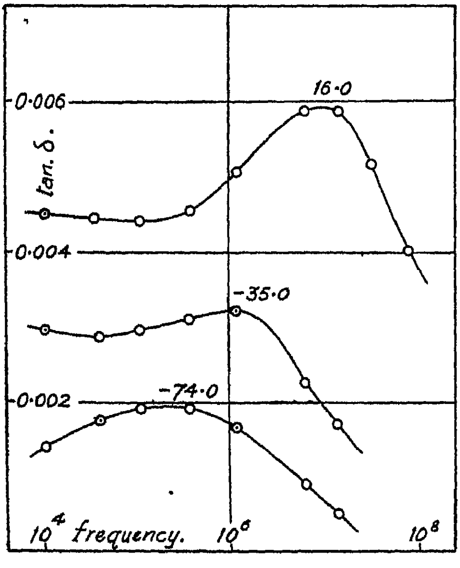


FIG. 3.—Power factor-frequency curves for 10 % dioctyl phthalate in polystyrene.

various esters in polythene " Alkathene " 20. The results are summarised in Table I.

TABLE I.

VALUES OF TAN δ_{max} AND f_{max} FOR POLYTHENE-ESTER MIXTURES AT CA. 16° C.

	Concentration.	tan δ_{max}.	f_{max}.
Dibutyl phthalate	0·6 %	0·0015	8×10^7
Dibutyl phthalate	2·0 %	0·0048	10×10^7
Dihexyl phthalate	2·0 %	0·0050	$3{\cdot}5 \times 10^7$
Dioctyl phthalate	2·0 %	0·0054	$1{\cdot}3 \times 10^7$
Tricresyl phosphate	0·06 %	0·0006	3×10^7
Tricresyl phosphate	2·0 %	0·0040	$1{\cdot}6 \times 10^7$
Propyl stearate	6·6 %	$>$0·003	$> 10 \times 10^7$

Discussion

(a) Distribution of Relaxation Times in Polyisobutene and Polythene Mixtures.—As pointed out above, all the curves obtained for polyisobutene and polythene mixtures indicate a distribution of relaxation times. Power factor-frequency curves for 2 % dioctyl phthalate in polyisobutene at 16·5° C. and in polythene at 17° C. and at −33·5° C., reduced to the same maximum tan δ and to the same value of f_{max}, are shown in Fig. 4. The results for the polyisobutene and polythene mixtures have been analysed in terms of the empirical equation due to Fuoss and Kirkwood:[10]

$$\tan \delta_f = A \operatorname{sech} (\alpha \log_e f_{max}/f),$$

where tan δ_f is the power factor at a frequency f, and α is a parameter which measures the breadth of the distribution of relaxation times. $\alpha = 1$ for a single relaxation time and is smaller as the distribution is broader. In the case of the polyisobutene and polythene mixtures the test plot gives a line composed of two straight portions changing slope at $f = f_{max}$, a consequence of the asymmetry of the tan δ-frequency curve. Values of α calculated from the test plot, and also average values obtained from the area under the tan δ-f curve and the value of tan δ_{max}

$$(\alpha = \pi \times \tan \delta_{max} / \text{Area})$$

are shown in Table II.

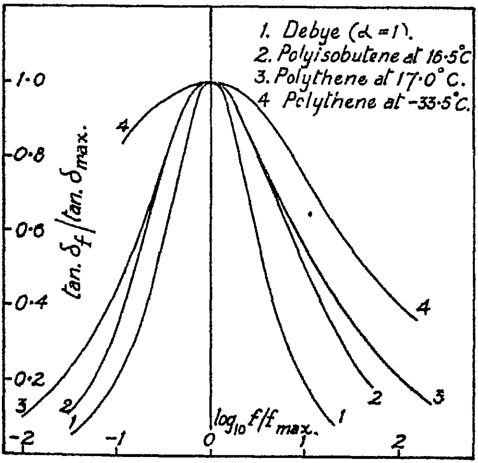


FIG. 4.—Relaxation time distribution. Comparative breadths of tan δ-f curves for polyisobutene and polythene containing 2 % dioctyl phthalate.

The decrease in α as the temperature is reduced, in the case of the amorphous polyisobutene, where there is no change of crystallinity, is similar to that observed in plasticised polyvinyl chloride by Fuoss and Kirkwood.[10] It seems probable that the lower values of α and the more marked decrease with a fall in temperature in the case of polythene is connected with the fact that, whereas polyisobutene is amorphous, polythene is partially (*ca.* 75 %) crystalline at the temperatures at which these measurements were made, and becomes less crystalline as the temperature

---

[10] Fuoss and Kirkwood, *J.A.C.S.*, 1941, **63**, 385.

is increased.[11] It is believed that each molecule of polythene passes through a number of crystallites, and hence there will be a variation in the ease of motion of the various segments of the chain. Those nearest a crystalline region will be less mobile, and hence will impart a greater restraint on the orientation of adjacent solute molecules than those in the middle of the less ordered and less densely packed amorphous regions.

TABLE II.—Relaxation Time Distributions.

Mixture.	Temp. °C.	Value of α.		
		Average.	High Frequency Side.	Low Frequency Side.
Polyisobutene —Dioctyl Phthalate	−2·5	0·61	—	—
	11·0	0·67	—	—
	16·5	0·67	—	—
	29·5	0·68	—	—
Polythene —Dioctyl Phthalate	−33·5	—	0·31	—
	−18·5	0·41	0·35	0·45
	−5·5	0·47	0·38	0·52
	17·0	0·57	0·46	0·65

(*b*) **Calculation of the Dipole Moment of the Solute.**—Sillars [3] has shown how the area under a tan δ-frequency curve (when the concentration of dipoles is low) is related to the concentration of dipoles and the dipole moment when there is a distribution of relaxation times. Using his equations, values for the dipole moment of dioctyl phthalate have been calculated from the area under the tan δ-frequency curves of Figs. 1 and 2, the curves being extrapolated where necessary as shown by the broken lines. Values of the dipole moment μ calculated in this way are shown in Table III. These values may be compared with figures given in the 1934 Faraday Society List for dimethyl phthalate (2·3-2·8) and diethyl phthalate (2·4-2·8).

TABLE III.—Values of Dipole Moment μ of Dioctyl Phthalate.

(*a*) In Polyisobutene.		(*b*) In Polythene.	
Temp. °C.	μ (Debye Units).	Temp. °C.	μ (Debye Units).
− 2·5	2·60	−33·5	2·60
11·0	2·54	−18·5	2·60
16·5	2·51	−5·5	2·60
29·5	2·46	17·0	2·54

(*c*) **Activation Energies of Dipole Orientation and Average Relaxation Times.**—The activation energy of dipole orientation ΔH_{DO} can be calculated approximately from the change with temperature of the average relaxation time τ_{DO}, which is given approximately by $2\pi f_{max} \tau_{DO} = 1$. Fig. 5 shows plots of $\log f_{max}$ against the reciprocal of the absolute temperature, based on the results in Figs. 1, 2 and 3. The polyisobutene points fall on a straight line, of slope corresponding to an activation energy of about 16 k.cal. The polythene points appear to lie on a curve, corresponding to a higher activation energy at lower temperatures; the average value over the range − 33·5 to + 17° C. is about 15 k.cal. The polystyrene points correspond to a much lower activation energy, about 5 k.cal.

The relation between the activation energies of dipole orientation and of viscous flow or high elastic deformation has been discussed by several authors. For polyisobutene, Kistler [12] gives a value of 24 k.cal. for an activation energy of viscous flow (although it seems that his method would be more closely concerned with high elastic deformation), and the viscosity data of Ferry and Parks [13] on a lower molecular weight (4,900 by a cryoscopic method) polyisobutene gives a figure of about 16 k.cal. at temperatures around 0° C. In the case of polythene, the activation energy of viscous flow of the molten material at 120-200° C. is about 10-12 k.cal.,[11] whereas work on the creep of the solid at 20-100° C., a predominantly elastic phenomenon, gives values of *ca.* 15-20 k.cal.[14] The activation energies of dipole orientation of a dissolved dipole will, of course, not be a function of the properties of the environment only, but the similarity in the values for the electrical and the mechanical activation energies seems to suggest that similar movements of the polymer chains are taking place in the cases of polythene and polyisobutene.

---

[11] Hunter and Oakes, *Trans. Faraday Soc.*, 1945, **41**, 49. Raine, Richards and Ryder, *ibid.*, 1945, **41**, 56, Bunn and Alcock, *ibid.*, 1945, **41**, 317.

[12] Kistler, *J. App. Physics*, 1940, **11**, 769.

[13] Ferry and Parks, *Physics*, 1935, **6**, 356.

In this connection, it may be noted that the activation energy calculated from the polythene-dioctyl phthalate results is lower than the value, about 28 k.cal., given by Jackson's results [8] on paraffin wax-ester mixtures, although it is close to the values obtained with similar mixtures by Sillars.[3] The relaxation times for these mixtures are appreciably greater than the values for the polythene mixtures at the same temperature, despite the much greater molecular weight and the greater viscosity in the molten state of the polythene. These differences can be explained by the assumption that in the paraffin wax experiments the ester molecules were rotating in a highly ordered crystalline solid, whereas in the polythene mixtures we believe that the ester molecules lie in the amorphous regions of the polythene.

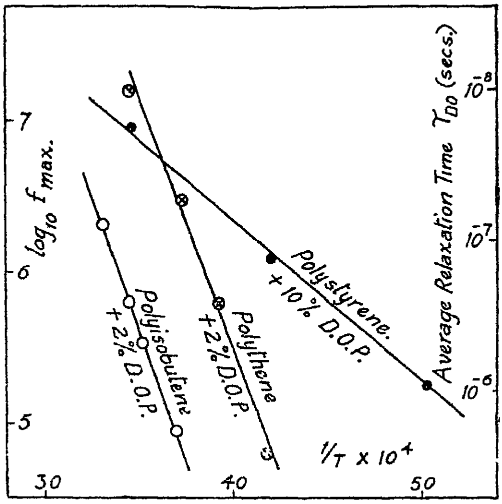


FIG. 5.—Effect of temperature on f_{max} for solutions of dioctyl phthalate.

Returning to the interrelation of the electrical and mechanical properties of polyisobutene and polythene, it may be noted that if we extrapolate the straight lines of Fig. 5 to a value of the electrical relaxation time ($\tau_{DO} = 1/2\pi f_{max}$) of about 1 second, we obtain temperatures of about −80° C. for polyisobutene and −100° C. for polythene. The brittleness temperatures of polyisobutene and of polythene (of high molecular weight), at which the relaxation time for high elasticity is of the order of one second, are about −70° C. and below −90° C. respectively. The similarity of the temperatures at which electrical and mechanical relaxation times have the same values has been pointed out for several polymers by Tuckett.[15]

The behaviour of the polystyrene mixtures is very different. The tan δ-f curves are not symmetrical and the areas under them are less than those under the curves for polyisobutene and polythene and, moreover, increase as the temperature rises. In addition, the activation energy is much less than the value of the activation energy of viscous flow of liquid polystyrene [16] (*ca.* 30-50 k.cal.). Similar subsidiary maxima (low temperature maxima) were observed by Fuoss[7] for plasticised polystyrene and polyvinyl chloride. In this case, it does not seem possible to identify the processes involved in dipole orientation and the mechanical deformation of the polymer. It may be suggested that the movement of the solute molecules in the electric field does not cause any twisting of the carbon-carbon skeleton of the polystyrene, but only causes minor movements such as the twisting of the phenyl groups relative to the rest of the polystyrene molecule.

---

[11] Oakes and Richards, *unpublished.*

[15] Tuckett, *Trans. Faraday Soc.*, 1944, **40**, 448.

[16] Based on data of Foote, *Ind. Eng. Chem.*, 1944, **36**, 245; and unpublished work by Oakes and Richards.

Finally, the effect of the nature of the solute molecule on the relaxation time may be considered briefly. Tan δ-f curves for modified polythenes, such as slightly oxidised polythene[17] or very lightly chlorinated polythene,[18] in which polar groups are attached chemically to the carbon-carbon skeleton, have a peak close to 10^8 cps. at ordinary temperatures (16-20° C.), indicating that the average relaxation time characteristic of the polyethylene chain is *ca.* 10^{-7} sec. It will be seen from Table I that some solutes lead to maxima at frequencies below 10^8 cps. when dissolved in polythene, others at around 10^8 cps. or at higher frequencies. The relaxation time will be governed in part by the activation energy of orientation of a solute molecule surrounded by sections of the polymer chains and the latter may be expected to be related in some way to the activation energies of movement, viscous or elastic, of sections of the polymer chains past each other and of the activation energy of viscous flow of the solute molecules. Of the limited range of solutes studied, the value of $f_{\max}$ decreases, *i.e.* the relaxation time increases, as the activation energies of viscous flow of the pure solutes increases (Table IV).

TABLE IV.—AVERAGE RELAXATION TIMES FOR DIFFERENT SOLUTES.

Solute.	ΔH_{visc} k.cal.	Average Relaxation Time (sec.) at 16-20° C.).
Propyl stearate .	4·0	$< 1{\cdot}3 \times 10^{-9}$
Dibutyl phthalate .	8·2	*ca.* $1{\cdot}3 \times 10^{-9}$
Dihexyl phthalate .	10·0	*ca.* $5{\cdot}0 \times 10^{-9}$
Dioctyl phthalate .	11·5	*ca.* $10{\cdot}0 \times 10^{-9}$
Tricresyl phosphate	15·0	*ca.* $11{\cdot}0 \times 10^{-9}$

As suggested by Fuoss,[7] further work on these lines with different solutes would be very instructive and would be a useful method of investigating the mechanism both of dielectric phenomena and also of viscous and elastic deformation.

Summary.

Tan δ-frequency curves have been obtained over a range of temperatures for solutions of esters in three contrasting hydrocarbon polymers—polyisobutene, polythene and polystyrene. All exhibit maxima at radio frequencies. The breadth of the curves for polyisobutene and polythene containing 2 % of dioctyl phthalate indicates a distribution of relaxation times; in the case of polythene the breadth of the distribution is the greater and also decreases more rapidly as temperature rises, due probably to the partly crystalline structure. The relaxation times are less for polythene than for polyisobutene at any temperature. From the areas under the curves, a value of 2·5-2·6 Debye units is calculated for the dipole moment of dioctyl phthalate. The activation energies of dipole orientation for dioctyl phthalate in polyisobutene and polythene are about 16 k.cal. and 15 k.cal. respectively.

The power factors of the polystyrene mixture are lower than that for the flexible polymers, and the activation energy calculated from the change of the power factor peak with temperature is only *ca.* 5 k.cal. It is believed that whereas dipole orientation in the flexible polymers involves

---

[17] Jackson and Forsyth, *J.I.E.E.*, 1945, 92, Pt. III, 23.

[18] Oakes and Richards, *unpublished.*

movements of the polymer chains similar to those in viscous or elastic deformation, in the rigid polystyrene there is a different mechanism.

We are indebted to Dr. P. A. Small of I.C.I. Ltd., Plastics Division, for viscosity data for the phthalates and tricresyl phosphate and also for his help in arranging for the preparation of specimens.

The authors wish to thank their colleagues at Manchester University and at Northwich for assistance and advice.

Dept. of Electro-Technics,
Manchester University.

Imperial Chemical Industries,
Research Department,
Alkali Division,
Northwich.

DIELECTRIC DISPERSION AND ABSORPTION IN NATURAL RUBBER, NEOPRENE, BUTAPRENE NM AND BUTAPRENE S, GUM, AND TREAD STOCKS.

By W. C. Carter, M. Magat,* W. C. Schneider and C. P. Smyth.

Received 11th March, 1946.

In order to increase our knowledge of the nature of rubbers, it was proposed that a dielectric study over a very wide frequency range be made on a variety of synthetic elastomers and natural rubber. The elastomers studied include vulcanised gum and tread stocks of natural rubber, of Neoprene, of Butaprene N as well as of Butaprene S elastomers differing in butadiene-styrene ratio and degree of conversion. In this paper, we shall consider only two of the Butaprene S rubbers, Butaprene S 85/15 medium conversion and Butaprene S 75/25 medium conversion. The measurements were carried out at three different temperatures, 20, 40 and 60° C. over a wide range of frequencies extending, in certain cases, from 60 cycles to 3000 megacycles.

Experimental.

The instruments and the experimental techniques employed in these determinations have been previously described.[1] However, for some tread stocks with high dielectric constants and high losses in the low frequency region, the method of calculating the dielectric constant had to be modified, because in these cases, the slightest non-parallelism of the electrodes plates becomes significant. Assuming the variable capacity of the condenser to be determined solely by its geometry, the value of the fixed capacitance C_f was determined for all the frequencies used and for different electrode distances. If C_2 and C_1 are the readings of the measuring condenser with and without the cell containing the sample, then the capacity of the sample is given by the relation

$$C_x = (C_1 - C_2) - C_f \quad . \quad . \quad . \quad . \quad (1)$$

and the dielectric constant, ϵ', is given by $\epsilon' = C_x/C_g$, in which C_g is the geometrical capacity of an air condenser of the same dimensions as those of the sample.

The samples selected for investigation were kindly prepared by Dr. R. F. Dunbrook and Dr. J. H. Dillon of the Firestone Tyre and Rubber Company. The recipes have been given previously.[2] Since the properties of rubbers depend on previous treatment, extra precaution was taken in making the samples uniform as to compounding and cure.

---

* Now at Laboratoire de Chimie Physique de la Sorbonne, 11 rue Pierre Curie Paris.

[1] Schneider, Carter, Magat and Smyth, *J. Amer. Chem. Soc.*, 1945, 67, 959.

[2] Tobolsky, Prettyman and Dillon, *J. App. Physics*, 1944, 15, 380.

Results and Discussion.

Gum Stocks. The dielectric data are recorded in the form of graphs (Figs. 1 and 2), and their most important features in Table I. The dielectric data for Neoprene have been given previously,[1] but will be referred to for the sake of comparison. The results are typical of plastic materials, *i.e.* the dielectric dispersion is quite broad, much broader than that predicted by the Debye theory, and the maximum losses are lower than those predicted by theory $(\epsilon''_{\max} = \frac{1}{2}(\epsilon_0 - \epsilon_\infty))$. In the low frequency range (1 kc.), the dielectric constants, ϵ', for Butaprene NM and Neoprene are high, 14·7 and 7·35 respectively because of their polar nature while the values for natural rubber and the Butaprene S rubbers are low, around 3·00, due to their non-polar structure. The decrease in ϵ' with increasing temperature is small. This can be attributed to the combination of three effects: the decrease in density, the increase in the kinetic energy of the moving segments leading to a greater randomness of motion and thus to a falling off of dipole orientation, and a loosening up of a rigid structure causing an increase in dipole orientation.[3]

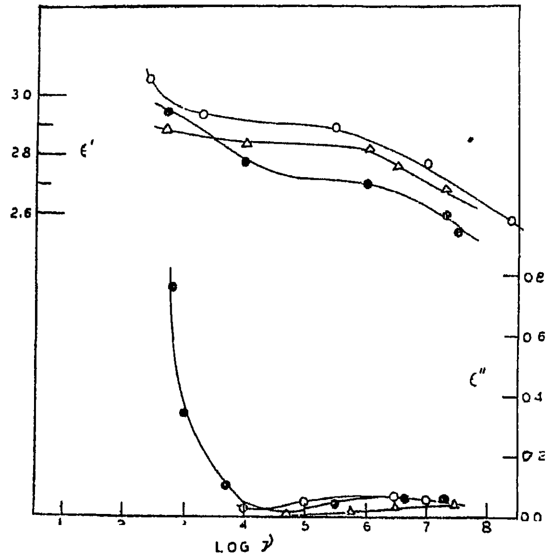


FIG. 1.—The dielectric constant and dielectric losses of Butaprene S 75/25 (gum) (R202) at various temperatures: ○ 20° C., ● 40° C., △ 60° C. As in the following graphs only a few representative points among the 20-30 measured are given.

As seen from Table I, column 4, the value of ϵ' at low frequencies, which is close to the static dielectric constant, decreases with increasing styrene content, as does also $\epsilon''_{\max}$ (column 5). The losses for polystyrene are so small that there is no maximum. For Neoprene[1] and Butaprene NM, the values are ten to fifteen times greater. The behaviour of the Butaprene S rubbers as well as natural rubber indicates the presence of groups containing dipoles which orient in an externally applied electric field. Those are carbon-sulphur bonds which give rise to a relatively small number of good-sized dipoles, and the conjugated or hyper-conjugated side chains which give rise to a large number of small dipoles; since the number of side chains decreases with increasing styrene content, they may be presumed responsible for the difference between polybutadiene and Butaprene S 75/25.

The sharp increase in the dielectric loss, ϵ'', at low frequencies when it occurs is attributed to direct current conductivity and is in no way connected with the relaxation process.

[3] Baker and Smyth, *J. Amer. Chem. Soc.*, 1939, **61**, 2063.

The critical frequencies, ν_c, given in column 8 are average values determined from the inflection point in the ϵ' — log frequency plots, the point of maximum loss, and the Cole and Cole plots,[4] in which the critical frequency is the frequency at which $\epsilon' = (\epsilon_0 - \epsilon_{\infty})/2 + \epsilon_{\infty}$, where ($\epsilon_0$ — the static dielectric constant, ϵ_{∞} = dielectric constant for infinite frequency). At 20° C. it is seen that the critical frequencies for all the samples are very nearly the same; one can hence expect that their inner viscosities are nearly equal. The critical frequency increase with temperature is presumably due to the decrease in the viscosity of the polymer. Butaprene NM has a lower critical frequency, 1·0 mc., and thus a more rigid structure or higher inner viscosity. Other Butaprene N rubbers for which no modifier was added showed still lower critical frequencies.

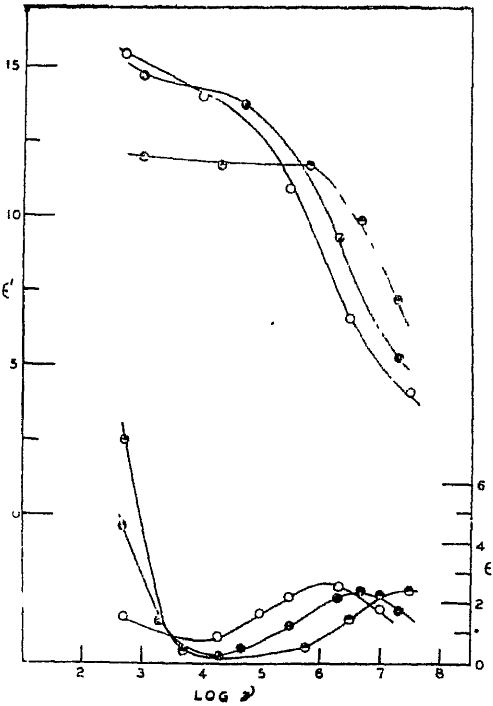


FIG. 2.—The dielectric constant and dielectric losses of Butaprene NM (gum) (R210), at various temperatures: ○ 20° C., ● 40° C., ◐ 60° C.

The values of $\epsilon_{\infty}^{\frac{1}{2}}$ as determined by the method of Cole and Cole[4] (see columns 6 and 7) tend to be slightly higher than the refractive index, as it was observed by Girard, Abadie and others, in many other cases.*

The orientation of polar units in an applied field requires a certain activation free energy, which may be determined from the critical frequency by means of the familiar rate equation,[5]

$$1/\tau_0 = \frac{kT}{h} e^{-\Delta F^{\ddagger}/RT} \quad (2)$$

in which $\tau_0 = 1/2\pi\nu_c$.

In column 8, the values of $\Delta F^{\ddagger}$ for the Butaprene S rubbers are seen to increase slightly with a temperature rise. The reverse is true for natural rubber, Butaprene NM and Neoprene.

From the free energy of activation at 20° C., it is possible to obtain an indication as to the length of the moving units giving rise to the relaxation. Assuming that the free energy of activation for dielectric relaxation is identical with that for viscous flow,[5] it is possible to calculate the energy of vaporisation, $\Delta E_{\text{vap.}}$, for 1 mole of the moving units by the use of the relation[6]

$$\Delta E_{\text{vap.}} = 2{\cdot}45\, \Delta F^{\ddagger}_{\text{vis.}} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

---

[4] Cole and Cole, *J. Chem. Physics*, 1941, 9, 341.

[5] Powell and Eyring, *Advances in Colloid Science*, Interscience Publishers, New York, N.Y., 1942, p. 217.

[6] Glasstone, Laidler and Eyring, *The Theory of Rate Processes*, McGraw-Hill Book Co., Inc., New York, N.Y., 1941, p. 492.

* One can hence expect to find a new dispersion region at wavelengths shorter than 1 cm. It will be remembered that Maar (*Z. Physik*, 1939, 113, 415), has found a dispersion region in ebonite with a maximum at 0·22 mm.

These "theoretical" values for $\Delta E_{\text{vap.}}$/mol. can be compared with the experimental values for $\Delta E_{\text{vap.}}$/cc. as determined from the swelling of the polymers in various solvents carried out on the same samples,[7] which, divided by ρ (density of the polymer) gives ΔE_{vap}/gr. By dividing

TABLE I.

Compound.	No.	t° C.	ε'_{1000}.	$\varepsilon''_{\max}$.	$\varepsilon^{\frac{1}{2}}_{\infty}$.	n (25° C.).	$\nu_{\text{crit.}}$ (Mc).	ΔF Kcal/mol.
Smoked sheet .	R198	20	3·01	0·13	1·62	1·52[(a)]	3·9±0·8	7·27
		40	2·94	—	1·59	—	11·2±3·0	7·13
		60	2·91	—	1·57	—	>30	—
Polybutadiene[(b)]	R214	20	3·06	0·19	—	—	—	—
		40	2·84	—	—	—	—	—
Butaprene S85/15	R200	20	3·01	0·09	1·60	1·52[(c)]	3·8±1·0	7·29
		40	2·84	0·08	1·59	—	6·0±2·7	7·52
		60	3·0	0·08	1·61	—	9·4±4·4	7·73
Butaprene S75/25	R202	20	2·95	0·07	1·63	1·53[(c)]	5·1±1·0	7·09
		40	2·82	0·07	1·59	—	7·0±2·7	7·42
		60	2·76	0·07	1·63	—	6·5±2·5	7·99
Polystyrene[(d)] .		25	2·64	0·0002	—	—	—	—
Butaprene NM .	R210	20	15·15	2·70	1·65	1·52[(a)]	1·03±0·02	8·10
		40	14·8	2·50	1·60	—	5·01±0·11	7·60
		60	12·0	2·45	1·63	—	16·8±0·7	7·32
Neoprene GN[(e)]	R216	20	7·48	0·97	1·89	1·56[(a)]	4·0±0·2	7·27
		40	7·18	1·00	1·76	—	26·1±0 7	6·59

(*a*) Wood, *Rubber Chem. Tech.*, 1940, **13**, 872. These values as well as those given under (*c*) refer to the unvulcanised polymer.

(*b*) Schneider, Carter, Magat and Smyth (unpublished)

(*c*) Data kindly supplied by Dr. B. L. Johnson, Firestone Tyre and Rubber Co. (unpublished).

(*d*) Yager, *Trans. Amer. Electr. Soc.*, 1938, **74**, 117.

(*e*) Schneider, Carter, Magat and Smyth, *J. Amer. Chem. Soc.*, 1945, **67**, 959.

TABLE II.

Sample No.	Natural Rubber R198.	Butaprene S 85/15 R200	Butaprene S 75/25 R202.	Noeprene GN R216.	Butaprene NM R210.
$\Delta E_{\text{vap. obs.}}$ (cal./cc.) .	70·1	72·2	73·1	84·5	92·1
ρ	0·985	0·985	1·014	1·27	1·03
$\Delta E_{\text{vap. obs.}}$ (cal./g) .	71·3	73·3	71·2	66·5	89·5
$\Delta E_{\text{vap. theoret.}}$					
(kcal./mole.) .	17·8	17·9	17·4	17·8	19·6
(g./mole.) . .	250	244	244	268	219
Base molec. weight/chain carbon atom .	17	15	17	22	17·8
Chain carbon atoms/in a segment	15	16	15	12 ‡	12
$\mu_{\text{calc.}}$ . . .	0·63	—	—	1·99	5·61
μ_{gas} . . .	0·38-0·68 [8, 9]	—	—	1·69-1·97 [9]	3·88-4·50 [10]

the energy of vaporisation per mole by the energy of vaporisation per gram, a value for the molecular weight of the moving unit is obtained. The simplest repeating unit of the polymer is known as the base mole. A quantity representing the distributed molecular weight per chain carbon atom is obtained by dividing the molecular weight of the base mole by the

---

[7] Gee, *Trans. Faraday Soc.*, 1942, **38**, 418; Scott, *Dissertation*, Princeton University, 1945 (*unpublished material*).

number of chain carbon atoms in the base mole.† From this and the molecular weight of the moving unit the number of carbon atoms moving as a unit can be found. A complete summary of the results is given in Table II. The number of carbon atoms in a segment moving as a unit varies from 12 to 16. Other workers [11, 12] have estimated that the moving segments contain 20 to 40 carbon atoms. We see that the order of magnitude is correct and that the length of the segments is nearly identical for all elastomers investigated. The absolute values depend naturally on the validity of the Eyring relation (3).

For all these rubbers, the plots of the function proposed by Fuoss and Kirkwood [13] against the logarithm of the frequency gave linearity in conformity with their theory. The slopes of these plots were used to calculate, by the method of Fuoss and Kirkwood, the dipole moment per monomer unit shown in Table II as μ_{calc}. These calculated values are not far from the measured values, listed as μ_{gas}, for the moment of small molecules similar to the monomer units. Each range of values given for μ_{gas} includes two or more molecules with different arrangements of methyl groups or of double bonds. In view of the small quantities of other substances present in the rubber as well as of the approximation necessary in the treatment, any better agreement between the calculated moments for the monomer units and the measured moments for the gas molecules would be fortuitous.

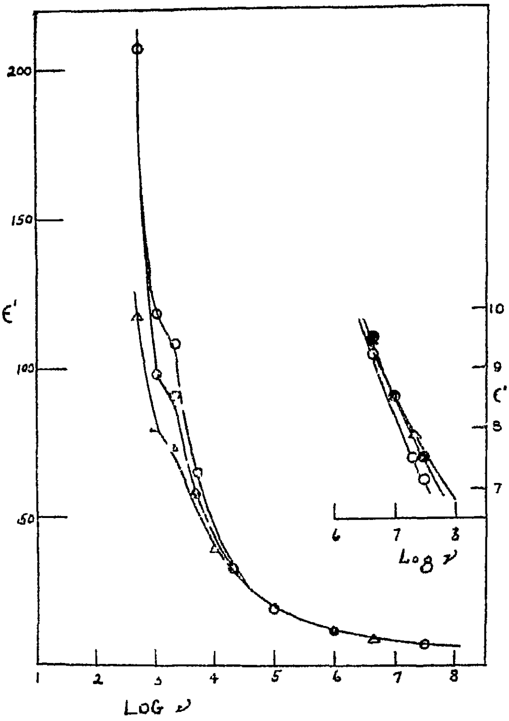


FIG. 3.—The dielectric constant of natural rubber tread (R199) at various temperatures : ○ 20° C., ● 40° C., △ 60° C. The results at higher frequencies are replotted on a very much larger scale in order to show the variation in ε' with increasing temperature.

Tread Stocks.—In Figs. 3-5 the dielectric constants and losses for the tread stocks of Butaprene NM and smoked sheet are given. For other non-polar elastomers the curves are similar to those for natural rubber.

In the case of Neoprene,[1] there is a loss maximum at 5Mc. at 20° C.,

---

[8] Hannay and Smyth, *J. Amer. Chem. Soc.*, 1943, **65**, 1931.

[9] Hannay and Smyth (*unpublished measurements*).

[10] Hurdis and Smyth, *J. Amer. Chem. Soc.*, 1943, **65**, 89.

[11] Flory, *J. Amer. Chem. Soc.*, 1940, **62**, 1057.

[12] Kauzmann and Eyring, *ibid.*, 1940, **62**, 3113.

[13] Fuoss and Kirkwood, *ibid.*, 1941, **63**, 385.

† The base molecular weights per chain carbon atom for copolymers were determined on the basis of the corresponding monomer.

‡ The number calculated for Neoprene GN is lower than the value 17 previously reported because of the use of our own value for $\Delta E_{vap.}$ instead of the one given by Gee.

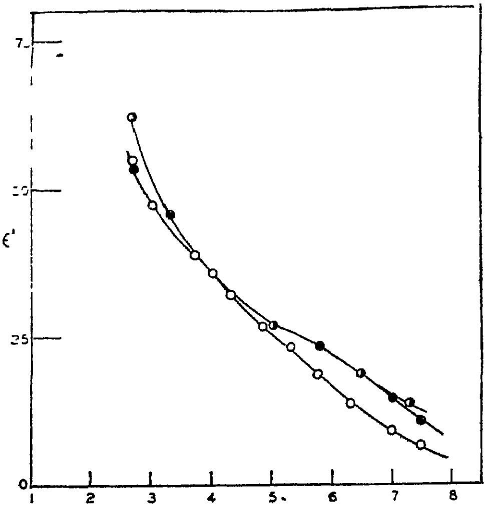


FIG. 4.—The dielectric constant of Butaprene NM (tread) (R211) at various temperatures: ⊙ 20° C., ● 40° C., ◑ 60° C.

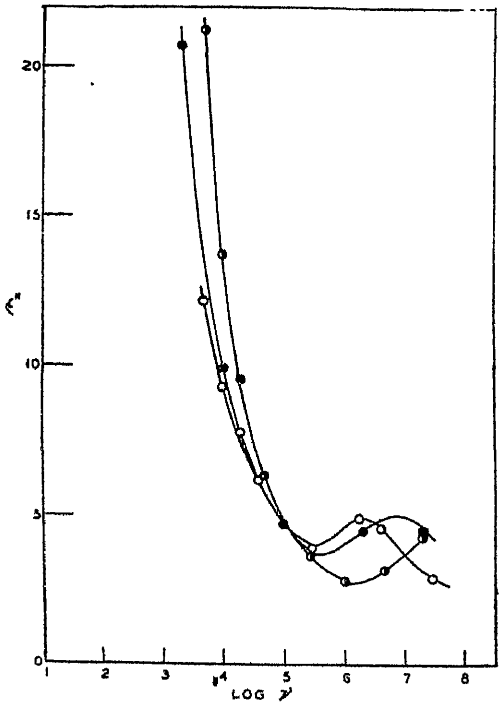


FIG. 5.—The dielectric losses in Butaprene NM (tread) (R211) at various temperatures: ⊙ 20° C., ● 40° C., ◑ 60° C.

which is associated with the Debye dispersion. In addition there is another maximum at 7 kc. quite independent of temperature. For non-polar elastomers, the small losses in the region of Debye dispersion observed (Fig. 1) for the gum stocks are completely obscured by the high losses due to the presence of carbon black. The presence of a maximum at 7 kc. in the loss curve for Neoprene as well as the presence of inflections in the $\epsilon' - \log \gamma$ curves for natural rubber (Fig. 3) and Butaprene S led us to an analysis of the data in terms of the Maxwell-Wagner [14] theory of absorption by composite dielectrics. According to the general theory of the dielectric behaviour of composite dielectrics, one has to expect in addition to the dispersion region, corresponding to the falling-off of orientation of the dipoles at higher frequencies (Debye), a second one corresponding to the so-called Maxwell-Wagner effect, the origin of which is an a.c. current which is in phase with the applied potential. This current, resulting from differences in the conductivities and dielectric constants of the substances composing the dielectric, introduces a time factor in the charging of the boundary surfaces. The formal equations relating ϵ', ϵ''_{MW} and τ_{MW} were discussed in details for Neoprene where the method used for the separation of the two effects was given.

Let it be said here that at each frequency the total loss observed is

$$\epsilon''_{obs.} = \epsilon''_{MW} + \epsilon''_{d.c.} + \epsilon''_{Debye}$$

and that the constants

necessary for the calculation of the Maxwell-Wagner losses, ϵ''_{MW} can be deduced from the real part of the dielectric constants in the same frequency region. The remaining losses consist of two parts. One part, ϵ''_{Debye} corresponds to pure Debye effect losses, while the low frequency portion obeys the law $\epsilon''_{d.c.} = 9 \times 10^{11} \times 4\pi\Lambda/\omega$ where Λ is the specific d.c. conductivity. Table III contains the values for $\epsilon_{\infty MW}$, and the specific d.c. conductivity, Λ, calculated both from the a.c. losses and from the d.c. conductivity determined directly in the cell by the application of Ohm's law to the direct currents flowing through the samples at different voltages ranging between 0 and 90 volts. Considering the difficulty of obtaining reproducible values for the dielectric properties of carbon black loaded stocks, the agreement can be considered as satisfactory (Fig. 6).

TABLE III.

	T (° C.).	ϵ_{∞} MW.	Λ mhos/cm. calcd. 20° C.	$\Lambda_{obs.}$ (27° C.).
Natural rubber (R199)	20	29·6	$1{\cdot}70 \times 10^{-6}$	$4{\cdot}75 \times 10^{-7}$
	40	29·0		
	60	28·9		
Butaprene S 85/15 (R 201)	20	22·2	$1{\cdot}22 \times 10^{-7}$	$1{\cdot}61 \times 10^{-7}$
	40	25·6		
	60	23·0		
Butaprene NM (R 211)	20	29·6	$1{\cdot}38 \times 10^{8}$	$2{\cdot}28 \times 10^{-9}$
	40	31·3		
	60	30·9		
Neoprene GN (R 217).	20	14·5	$8{\cdot}1 \times 10^{-10}$	$5{\cdot}9 \times 10^{-10}$
	40	13·6		
	60	13·8		

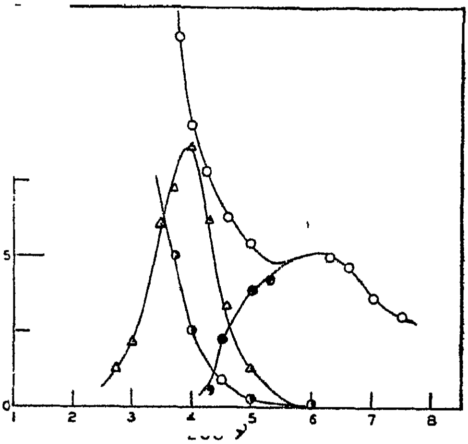


FIG. 6.—The Maxwell-Wagner analysis of dielectric losses in Butaprene NM (tread) at 20° C.

○ $\epsilon''_{obs.}$; △ ϵ''_{MW}, ◑ $\epsilon''_{d.c.}$,
● $\epsilon''_{Debye} = \epsilon''_{obs.} - [\epsilon''_{MW} + \epsilon'_{d.c.}]$.

As far as the Debye dispersion is concerned, it is interesting to note that the critical frequencies are nearly identical for the gum and the tread stocks of Neoprene (Table IV). Carbon black is known to reinforce Neoprene only slightly. The negligible shift in ν_c suggests that there is no interaction between carbon black and Neoprene.* For natural rubber and Butaprene S 85/15 treads, there seems to be a definite shift in the

---

[14] See Whitehead, *Lectures on Dielectric Theory and Insulation*, McGraw-Hill Book Co., Inc., New York, N.Y., 1927, p. 99.

* Further evidence has been given in the paper on Neoprene already referred to, in which it was shown that the plots of the Fuoss-Kirkwood function against the logarithm of the frequency at 20° C. are identical for the gums and tread stocks.

critical frequency with the addition of carbon black (Table IV), suggesting strong interaction with the carbon black in these elastomers. The small shift in the critical frequency for Butaprene NM indicates interaction intermediate between that for the non-polar rubber and that for Neoprene.

The effect of carbon black upon the dielectric properties of rubbers as clearly shown by the analysis of the dielectric losses of the tread stocks is perfectly in line with the mechanical reinforcement as observed on the same samples.

TABLE IV.

Sample.	No.	ν_c (mc) 20° C.	$\varepsilon''/\varepsilon'$. 20° C.
Natural rubber (gum) .	R198	3·9	0·047
(tread) .	R199	0·2	0·057
Butaprene S85/15 (gum)	R200	3·8	0·034
(tread)	R201	0·1	0·033
Neoprene GN (gum) .	R216	4·0	0·18
(tread) .	R217	5·0	0·17
Butaprene NM (gum) .	R210	1·0	0·30
(tread) .	R211	1·5	0·37

* Rough estimates by Maxwell-Wagner analysis.

Summary.

Dielectric constants of various elastomers, gums and treads, were measured over a wide range of frequencies. The data are well represented by the Fuoss-Kirkwood theory which leads to reasonable values for the dipole moments involved. A comparison of the Debye critical frequencies with the "heats of vaporisation" deduced from swelling experiments shows that the movable segments are all of the same size. The abnormally high losses of tread stocks are due to a superposition of Debye effect, Maxwell-Wagner effect and d.c. conductance.

The writers wish to express their indebtedness to the Firestone Tyre and Rubber Company for their support of the work and their co-operation in supplying the samples and to the Rubber Reserve Company for additional assistance.

Frick Chemical Laboratory,
Princeton, New Jersey.

GENERAL DISCUSSION

Dr. H. Fröhlich (*Bristol*) said: It might be interesting to measure the relaxation times of stretched material.

Dr. A. Schallamach (*Welwyn Garden City*) said: The calculation of the size of the moving units is based on the equality of the free energies of activation for dielectric relaxation and viscous flow, and on the empirical relation, the authors' eqn. (3).* It can be shown, however, that even in cases where there is strong reason to admit this equality these assumptions lead sometimes to absurd values of the molal volume in question. From the well-known equation

$$\eta = (hN/V)e^{\Delta F^{+}/RT}$$

and the authors' eqn. (2)* one obtains, ΔF^{+} assumed to be the same in both cases, the relation

$$V = RT\tau/\eta$$

Taking here, for example, a few compounds of a chain length comparable with those found by the authors for their moving segments, one has [1]

geraniol (213° K.) V = 640·0 cc.
phytol (229° K.) V = 61·3 cc.
di-dihydrocitronellyl ether (200° K.) V = 26·0 cc.

---

[1] Schallamach, *Trans. Faraday Soc.*, 1946, **42**, 495.
* Page 215.

The disturbing fact is not so much that the absolute values are wrong but that the ratio η/τ increases with rising molecular weight as from geraniol to phytol and does not decrease as suggested by the above relation for V. The failure of the viscosity equation proposed by Eyring is most probably due to the model on which it is based.

It appears to be a generally observed experimental fact that, with the same solute in solvents of different viscosity or in pure liquids of different molecular weights, changes of the dielectric relaxation time are far less pronounced than those of the viscosity.[1, 2, 3, 4]

Dr. R. F. Tuckett (*Welwyn*) said : I should like to add a little to Dr. Schallamach's remarks on the connection between relaxation time and viscosity in long chain molecules—the discrepancies found in medium length compounds are still more marked in polymers. From Flory[5] and Eyring's[6] work, the viscosity (η) of a polymer is related to its molecular weight, M, by an expression of the form

$$\log \eta = A + B\sqrt{M}$$

whereas the few results on the dependency of electrical relaxation time (τ) on molecular size suggest a linear relationship between τ and M. The theoretical reasons for this are by no means clear—Kirkwood and Fuoss'[7] theory is in agreement with experiment but it is only concerned with a single chain completely isolated from its neighbours—this theory, therefore, is that of a very dilute solution of a polymer in an inert solvent and cannot be applied to the measured polymer-plasticiser systems in which there must be strong interaction and interference between neighbouring chains.

On the connection between elastic and electrical properties of polymers, it is essential to choose the correct absorption bands when making the comparison—it is important to distinguish between dipole rotations caused by movements of the main backbone carbon and those rotations which can occur independently of the main chain[8]—this is best seen in polymethyl methacrylate where two absorption bands occur in different temperature regions at the same frequency.[9, 10] In this connection, I would like to ask if there exists any convincing mechanism which explains the very broad and shallow absorption bands which occur in polyvinyl chloride[11] and other polymers[12] at low temperatures—the bands at higher temperatures seem explicable on the basis of segmental rotations of the main chain as in highly elastic extension.[8] The cause of the low temperature bands, in the region where the material is a rigid solid, is still obscure.

Mr. L. Mullins (*Croydon*) (*communicated*) : Richards and Magat have stressed that the moving units which take part in the readjustments during mechanical deformation and electrical displacement are of the same size. Typical measurements, which we have made on a number of soft vulcanised rubbers, of the change of mechanical energy absorption with temperature are shown in Fig. 1.

These curves resemble the familiar power-factor/temperature curves (Fig. 2), possessing a maximum energy absorption at a temperature characteristic of the material ; furthermore, electrical measurements made at a similar frequency have a maximum dielectric loss at similar temperatures.

[2] Martin, *Physik. Z.*, 1936, **37**, 665.
[3] Jackson and Powell, 1946, **42A**, 101.
[4] Whiffen and Thompson, 1946, **42A**, 114.
[5] Flory *J.A.C.S.*, 1940, **62**, 1057.
[6] Kauzmann and Eyring, *J.A.C.S.*, 1940, **62**, 3113.
[7] Kirkwood and Fuoss, *J. Chem. Physics*, 1941, **9**, 329.
[8] Tuckett, *Trans. Faraday Soc.*, 1944, **40**, 448.
[9] Mead and Fuoss, *J.A.C.S.*, 1942, **64**, 2389.
[10] Ponomarev, *J. Tech. Phys. U.R.S.S.*, 1940, **10**, 587.
[11] Fuoss, *J.A.C.S.*, 1941, **63**, 369.
[12] Mead and Fuoss, *ibid.*, 2832.

Values of the apparent energy of activation of the processes taking place in both mechanical and electrical deformation, calculated from this temperature dependence, compare reasonably well, and are similar to the values reported for the viscous flow of unvulcanised rubbers. However, although these results suggest a similar relaxation mechanism, the parallelisms between electrical and mechanical energy absorption are not complete; all rubbers exhibit a temperature at which there is a maximum and very large mechanical energy absorption, whereas non-polar rubbers show little dielectric loss over the whole temperature range (Fig. 2).

Contrary to the findings of Magat in his electrical measurements the introduction of carbon black into natural rubber vulcanisates causes no definite shift of the temperature of maximum mechanical energy absorption. (A fuller account of this work has now been published.)[13]

Mr. R. B. Richards (*Northwich*) (*communicated*): One must be careful when carrying out a quantitative resolution of loss-factor frequency curves for mixtures containing carbon black (such as tread stocks) into contributions due to d.c. conductivity, dipole rotation and "Maxwell-Wagner" loss in the manner described by Carter, Magat, Schneider and Smyth. The form of the loss-factor frequency curve for the Maxwell-Wagner contribution used in their paper was developed for dispersions of spheres. It has been suggested[14] that carbon black particles dispersed in rubber form aggregates of chain-like form and that these are the cause of the high conductivity of carbon black filled rubbers. If such chain-like aggregates are present, Wagner's equation for loss-factor as a function of frequency is no longer valid, and Sillars[15] has shown that the loss-factor of a dispersion of elongated particles is higher and bears a different relation to frequency than that of a dispersion of an equal volume of spheres of the same material. We have found that whereas the Maxwell-Wagner equations represent very well the loss-factor frequency curve for dispersions of water droplets in polythene they are quite inadequate, in the form used by Carter, Magat, Schneider and Smyth, to describe the dielectric properties of dispersions of carbon black in the same dielectric. The results of Kemp and Herrmann on the effect of additions of carbon black to rubber and "Paragutta" also seem incompatible with the simple Wagner equation, which would lead to a linear increase of loss-factor with the content of the disperse phase.

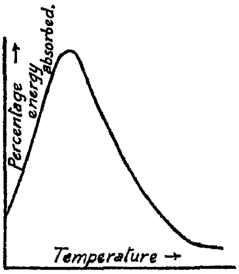


FIG. 1.

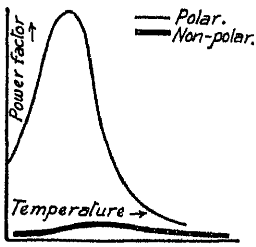
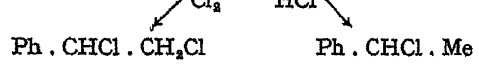


FIG. 2.

Dr. R. J. W. Le Fèvre (*Farnborough*) said: The assumption that during the early stages of the chlorination of polythene, all hydrogen atoms are equally likely to be replaced by chlorine, is unsafe if two or more mechanisms are simultaneously operating. The example of ethyl benzene can be quoted as a warning. This hydrocarbon yields, by direct halogenation, a mixture of Ph . CHCl . Me and CH_2Ph . CH_2Cl plus polychloroderivatives—the first often preponderating. Yet evidence exists[16] that the *first* attack is at the β-position, intermediate olefine formation occurs, followed by addition either of Cl_2 or HCl :—



$$\begin{array}{c}
Ph\,.\,CH_2CH_3 \rightarrow Ph\,.\,CH_2CH_2Cl \rightarrow Ph\,.\,CH:CH_2 \\
\swarrow Cl_2 \qquad HCl \searrow \\
Ph\,.\,CHCl\,.\,CH_2Cl \qquad\qquad Ph\,.\,CHCl\,.\,Me
\end{array}$$

With polythene, therefore, as soon as reaction has taken place at any C atom, the immediate neighbours of this atom have a higher chance than those more distant of being the next to receive Cl substituents.

Mr. R. B. Richards (*Northwich*) (*partly communicated*) : Mr. Garton has suggested that the results quoted by Plessner and Richards for solutions of esters in polystyrene indicated that the materials under test were dispersions rather than true molecular solutions. All of the samples were glass-clear and there are no signs of "sweating-out" of the esters several months after their preparation, even in the case of mixtures containing higher ester contents than those for which results are quoted. This alone would not, of course, disprove the presence of small local aggregates of ester molecules such as those discussed by Dr. Schallamach. It is possible that a study of changes with temperature and concentration of the dielectric properties of two-component polymer systems might throw much light on the formation of local aggregates of the separate components at the temperature or concentration at which visible separation into two phases occurs, and hence on such problems as the mechanism of plasticizer action.

With regard to Dr. Le Fèvre's comments on randomness of chlorination, I agree that it is probable that even in the earliest stages of chlorination the chlorine atoms will tend to cluster together in groups along the polymer chain. The figures for the average dipole moments of the C—Cl bond in the chlorinated polythenes of lowest chlorine content (2 to 25 %) suggest a considerable measure of dipole interaction, whereas if chlorination were entirely random this would only become marked at higher chlorine contents. Thus a statistical calculation of the proportion of chlorine atoms in various types of group at various total chlorine contents (for which I am indebted to Mr. W. F. L. Dick of this laboratory) shows that at a total chlorine content of 8 % by weight, less than 2 % of the chlorine atoms would be in $—CCl_2—$ groups or in adjacent —CHCl— groups and this proportion would only reach 10 % at a chlorine content of about 35 % by weight.

Since the paper on the dielectric properties of chlorinated polythenes was written, Dr. H. W. Thompson has compared the infra-red spectra of two pairs of chlorinated polythenes of the same chlorine contents, 8 % and 40 %, each pair containing a hot and a cold chlorination product, and has found, as we suggested on the basis of the average dipole moments per C—Cl bond, that the concentration of $—CCl_2—$ groups relative to that of —CHCl— groups is in each case higher in the case of the cold chlorinated sample.

Dr. H. Fröhlich (*Bristol*) said : It has been suggested that a sudden strong increase of dielectric constant indicates that beyond the temperature at which this occurs dipoles or whole molecules should be free to rotate. I should like to mention that this need not be so. The same behaviour will be found if a dipole has two or more equilibrium directions with the same energy. In the case of long chains I rather think that the molecules do not rotate but have two equilibrium positions.

---

[13] *J. Rubber Res.*, 1946, **15**, 60.
[14] Kemp and Herrmann, *Proc. Rubber Tech. Conf.*, 1938, p. 893.
[15] Sillars, *J.I.E.E.*, 1937, **80**, 378.
[16] *J. Chem. Soc.*, 1927, 1159.

G. SOLIDS.

CRYSTAL STRUCTURE OF BARIUM TITANIUM OXIDE AND OTHER DOUBLE OXIDES OF THE PEROVSKITE TYPE.

By Helen D. Megaw.

Received 20*th March*, 1946.

The exceptionally high dielectric constant of barium titanium oxide has attracted much attention to this compound in recent years. The present investigation of this and related compounds was undertaken as a means of clearing the ground on which a theoretical treatment of the dielectric properties might be built.

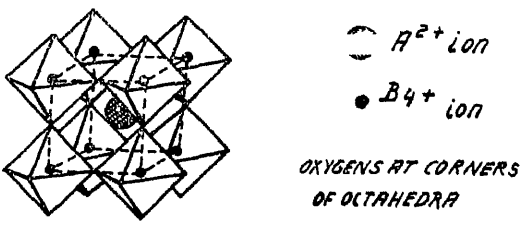


Fig. 1.—Ideal perovskite structure.

All the compounds investigated have the empirical formula

$$A^{2+}B^{4+}O_3,$$

and belong to the perovskite group of structures. These are all closely related to the ideal perovskite type, which has a simple cubic cell with the formula number of atoms per cell ($Z = 1$) and atomic parameters as follows:

Small cation, B^{4+} . .	000
Large cation, A^{2+} . .	$\frac{1}{2}\frac{1}{2}\frac{1}{2}$
Oxygens	$\frac{1}{2}00$, $0\frac{1}{2}0$, $00\frac{1}{2}$.

(It is equally legitimate, and sometimes done, to take the origin at the point occupied by the large cation, in which case all the atomic parameters must be increased by $\frac{1}{2}\frac{1}{2}\frac{1}{2}$.)

The resultant structure is shown in Fig. 1. It consists of a network of regular BO_6 octahedra linked by their corners throughout space, and holding the large cations in the holes between the octahedra where they are surrounded by 12 oxygens. Alternatively, it may be thought of as derived from a cubic close-packing of oxygens in which one out of four oxygens is removed and replaced by the large cation, the small cation occupying the octahedral hole between the oxygens.

An ionic structure, as Goldschmidt[1] originally showed, is normally only stable when all the ions are in contact so that none of them is able to "rattle" in a hole enclosed by ions of opposite sign. In the perovskite structure, if all the ions are thought of as rigid spheres, this is only true if the geometrical condition holds,

$$R_A + R_O = \sqrt{2}(R_B + R_O),$$

where R_A, R_B, R_O are the ionic radii of the large cations, small cations, and oxygens respectively. Since in fact the ions are to some extent compressible, Goldschmidt defined a tolerance factor t from the relation:

$$R_A + R_O = t\sqrt{2}(R_B + R_O),$$

---

[1] Goldschmidt, *Geochem. Vert. Elem.* VII, VIII, 1927.

where the radii are those of any atoms taking part in a compound ABO_3 that might possibly have the perovskite structure. He found empirically that t lay between 1 and a lower limit of about 0·80, below which structures other than the perovskite one were formed. The conditions he put forward for the occurrence of the structure may be expressed as follows:

(1) the smaller cation must be capable of 6-coordination with oxygen, in order to form the octahedron network;

(2) the larger cation must be nearly large enough to fill the holes between the octahedra.

The first condition is always satisfied in the compounds with which we have to deal, but the second requires further consideration. If the large cation is appreciably too small, the perovskite structure is not formed, as Goldschmidt showed. But more recent work [2, 3, 4] has established that the size limits within which the structure is truly cubic are narrower than those set by him, and that less symmetrical variants of the structure may occur with t value greater than 1 as well as in the lower part of his permissible range, the modifications at the two ends of the range being of different type.

Table I shows the values of t for the compounds considered in this investigation. In calculating this, the Goldschmidt radii have been used, except for Zr^{4+} where Naray-Szabo [5] has shown that an amended value of 0·77 A. is required (and this is substantiated by the present work). The A radii were corrected by the factor appropriate for 12-coordination. Figures in brackets, included for comparison, refer to compounds which either do not exist or were not examined.

TABLE I.

B A	Ti	Sn	Zr	Th
Ca	0·89	0·85	0·84	(0·72)
Sr	0·97	0·92	0·91	(0·78)
Ba	1·02	0·97	0·96	0·83
Pb	0·98	(0·92)	0·93	(0·79)
Cd	0·88	(0·83)	(0·82)	(0·71)

All the compounds here have a structure of the perovskite type, but not all have the ideal cubic form of it which has so far been considered. The cubic structures have been enclosed by a framework in the table. They all have values of t lying between 0·97 and 0·91 with the exception of $BaThO_3$, and in this case there seems reason to doubt the accepted value of the Th^{4+} radius.

The four compounds with a lower value of t, namely $CaTiO_3$, $CaSnO_3$, $CaZrO_3$, $CdTiO_3$, all show a distorted perovskite structure, with lower symmetry and a larger unit cell. This has been known in a general way from the early work, but the details have only recently been established and are still not complete. The type of distortion is the same in all four; it is just what would be expected on steric grounds in order to bring oxygens into contact with the large cation A, while leaving the distance B—O unaltered. It consists in effect in a very small shear of the cubic lattice so as to make it monoclinic, accompanied by a slight change of length of the b axis (normal to the plane of the shear) relative to the a and c axes (lying in the plane) which, however, remain equal to each other. The experimental evidence for this [6] is based on very exact determinations of the cell constants from the high-angle lines of powder photographs taken with a 19 cm. camera. A consequence of this equality is that the lattice may be referred to rectangular axes, the new a and c axes being

---

[2] Hoffmann, *Z. physik. Chem.*, *B*, 1935, **28**, 65.
[3] Megaw, *Nature*, 1945, **155**, 484.
[4] Naray-Szabo, *Naturwiss.*, 1943, **31**, 203.
[5] Naray-Szabo, *ibid.*, 1943, **31**, 466.
[6] Megaw, *Proc. Physic. Soc.*, 1946.

chosen to bisect the angles between the original *a* and *c*. Thus the true symmetry is orthorhombic. However, for the purposes of the present paper it is very much more convenient to continue to use the original monoclinic axes, since their derivation from the cube axes is so direct, and since this description of the structure best brings out the resemblance to all the related structures.

The distortion of the original lattice to monoclinic is accompanied by changes in the atomic parameters which make it necessary to double the cell edge before a repeat unit is obtained. The present work confirms the existence of such a doubled cell, but does not attempt to determine the atomic parameters. This has been done by Naray-Szabo[5] for $CaTiO_3$, though his results will need slight amendment in view of the true orthorhombic symmetry. He found that only the oxygen ions were displaced from the positions occupied by them in the cubic structure; the cations remained at their original positions. Such evidence as can be drawn from the present work supports this conclusion and extends it to the other Ca compounds; but for $CdTiO_3$ it appears that the Cd ion is displaced so as to bring it into a position more resembling 6-coordination.

It seems clear that the departure from cubic symmetry in these orthorhombic compounds is due to steric considerations only, and can be completely accounted for from that point of view.

The three remaining compounds in Table I, namely, $BaTiO_3$, $PbTiO_3$, and $PbZrO_3$, all have fairly large *t* values and are all tetragonal. Here the cubic lattice is deformed by a uniform extension or compression along one tetrad axis, but the atomic parameters are unchanged and there is no doubling of the cell edge (unless possibly in $PbZrO_3$, which has been less fully examined). The differences in the axial ratios are noteworthy, as shown in Table II.

TABLE II.

AXIAL RATIO *c*/*a* FOR TETRAGONAL STRUCTURES.

$BaTiO_3$ . .	1·0100 ± 0·0001
$PbTiO_3$ . .	1·0635 ± 0·001
$PbZrO_3$ . .	0·988 ± 0·001

It is difficult to see how purely steric considerations can explain these structures. In order to examine this point more closely, we must consider the interatomic distances and compare them with the sum of the Goldschmidt radii. These are shown in Table III for the three compounds in question, and values from the cubic compounds are included for purposes of comparison.

In the two lead compounds, the Pb^{2+} cation is not too large for the available hole, and there seems no need for the distortion. The explanation probably lies in the structure, as distinct from the size, of the Pb^{2+} ion, which is large and highly polarisable; it cannot be treated, except very roughly, as an undeformable sphere, and it seems likely that it forms directed bonds. No detailed account of this can be put forward without further experimental evidence, and there is no explanation why the axial ratio is greater than unity in $PbTiO_3$ and less in $PbZrO_3$.

The case of $BaTiO_3$ is more striking, and as this is the compound of most interest for its high dielectric constant it has been examined in considerable detail. The Ba ion is too large for the hole in the cubic structure (as shown in Table III, cols. 5 and 6), and it pushes the oxygens apart, leaving the Ti's between them free to "rattle". This "rattling" is not cured by the tetragonal distortion, for all the Ti—O distances are still too big (as can be seen from Table III, cols. 3 and 4) though not all equally so. Fig. 2 shows this in section, all the ions being drawn to scale with their Goldschmidt radii.

At this point it is necessary to consider whether the Ti ion may be displaced from the lattice corners: that is, to reopen the question whether the atomic parameters are accurately the same as in the cubic structure. That they are approximately so follows from the general agreement of calculated and observed intensities of X-ray reflections; but the intensities have not been measured accurately enough to say whether or not very small displacements occur. Such displacements must in any case be consistent with the symmetry of the small tetragonal unit cell, since there is no evidence of a doubled cell as in $CaTiO_3$; and in practice this limits them to displacements parallel to the tetrad axis. There are then two possibilities; either the displacements of the Ti's are distributed at random in the positive and negative directions of the axis, or they are all in the same direction, and the crystal loses its centre of symmetry.

TABLE III.

INTERATOMIC DISTANCES AND SUM OF GOLDSCHMIDT RADII.

	Compound.	B—O.	$R_B + R_O$.	A—O.	$R_A + R_O$.
Ti—O . .	$SrTiO_3$	1·949	1·96	2·755	2·69
	$BaTiO_3$	1·992 2·013		2·818 2·832	2·84
	$PbTiO_3$	1·948 2·072		2·755 2·844	2·72
Sn—O . .	$SrSnO_3$	2·012	2·06	2·845	2·69
	$BaSnO_3$	2·054		2·907	2·84
Zr—O . .	$SrZrO_3$	2·047	2·09	2·894	2·69
	$BaZrO_3$	2·091		2·957	2·84
	$PbZrO_3$	2·075 2·050		2·934 2·917	2·72

Even if these displacements occur, it is clear that ionic central forces are not adequate to explain the environment of the Ti atom. The Ti is closer to four or possibly five of its neighbours than to the remaining one or two. We have seen above that steric requirements cannot explain this, and hence it seems necessary to postulate a set of directed bonds, stronger in the (001) plane than perpendicular to it, superimposed on the ionic forces which are predominant in determining the structure. These directed bonds cannot reasonably be attributed to the Ba, and must have their origin in the Ti—O link. There is no previous evidence from any other compound of such a planar set of bonds belonging to Ti (though Pauling considers there is an octahedral set in TiS_2); but it is suggested here that the abnormal volume available to the Ti^{4+} ion in this structure has made possible a bond

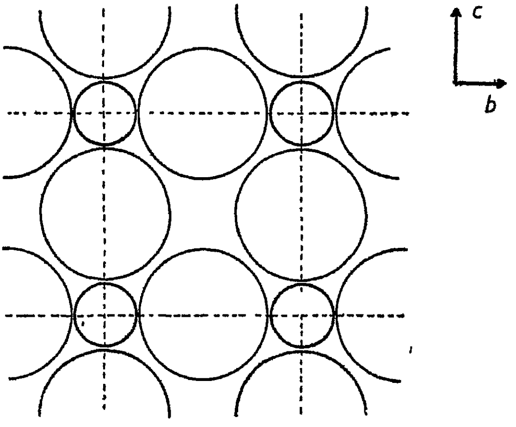


FIG. 2.—Section of structure of $BaTiO_3$ in the plane (100) at height $x = 0$. Large circles represent oxygen, small circles titanium (both drawn to scale with their Goldschmidt radii).

system which would not be stable otherwise. The energy states of ions are well known to be sensitive to volume changes.

When $BaTiO_3$ is heated, it passes through a transition point in the neighbourhood of 120° C. This is a λ-point, and the transition is clearly analogous to the α—β transition in quartz, as will be seen later. The high-temperature form is cubic, and has the ideal perovskite structure. It is suggested as an explanation of this transition that the directed bond system disappears, leaving a purely ionic structure at the higher temperatures. The change may be brought about by the volume changes associated with thermal expansion.

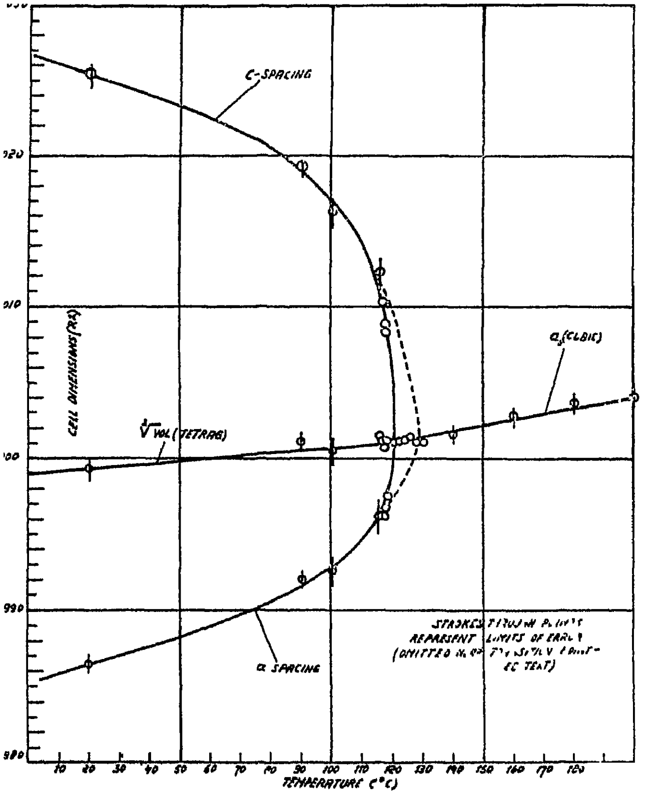


FIG. 3.—Variation of spacing of $BaTiO_3$ with temperature.

The changes in the cell dimensions have been followed in detail throughout the transition. They are continuous (at least within the limits of experimental error). Fig. 3 shows the variation of the *a* and *c* dimensions with temperature ; (the cube-root of the volume of the tetragonal form is included in the same diagram). The shape of this curve is characteristic of a λ-point change, and if we calculate from it the thermal expansion coefficients, shown in Fig. 4, this fact is still more obvious. These latter, of course, are only very approximate, but they clearly have the general λ shape. They should be compared with the corresponding α-β transition of quartz, as determined by Jay[7].

Whether the curves of Fig. 3 are discontinuous in the immediate neighbourhood of the transition point, or have an abrupt change of direction, or curve round rapidly but continuously, cannot be determined from the experimental data available, and is perhaps meaningless to ask. This, of course, applies still more emphatically to the detail of the peak or discontinuity in the thermal expansions derived from them.

---

[7] Jay, *Proc. Roy. Soc.*, 1933, **142**, 237.

There are certain anomalous intensity changes near the transition point which cannot be discussed fully here. Two separate effects may be distinguished: (1) the coexistence of cubic and tetragonal forms, in amounts dependent on the temperature, over a range near the transition point; (2) imperfections of the crystal lattices near the limits of their ranges of stability. The coexistence of the two forms is comparable with that observed in NH_4Cl, which has been explained very neatly by Dinichert and Weigle[8]. It is accounted for by the stresses in the lattice set up by the transition process itself.

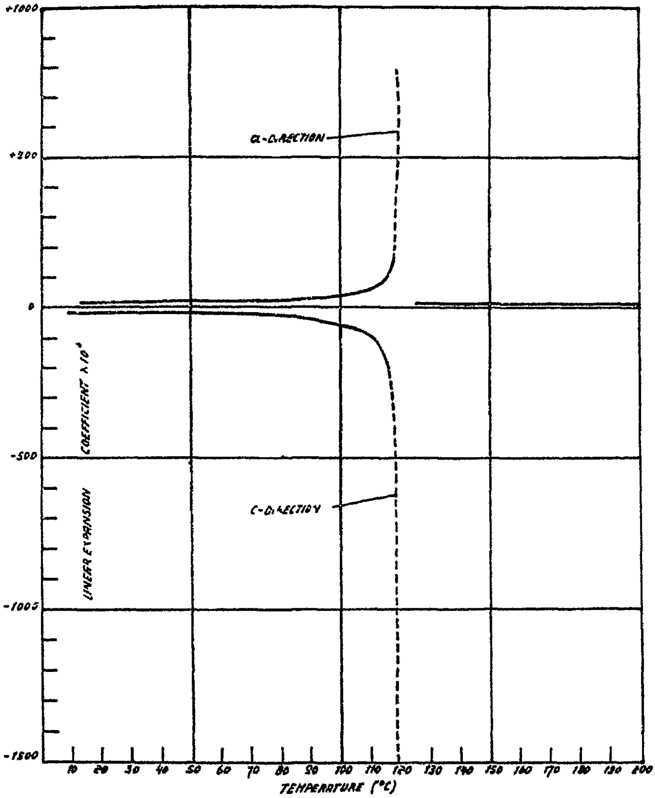


FIG. 4.—Linear thermal coefficients of $BaTiO_3$.

The continuous change of cell dimensions at the 120° C. transition point is, of course, accompanied by a continuous change of axial ratio, which decreases from 1·0100 to 1 between room temperature and the transition point. Photographs taken *below* room temperature, however, also show a decrease of axial ratio, as indicated in Fig. 5. Unfortunately it was not possible to make accurate determinations of the cell dimensions at these temperatures, partly owing to the apparatus available, but mainly to the anomalous intensity changes which obscured the clarity of the

[8] Dinichert and Weigle, *C.R. Soc. Phys. Geneve*, 1944, 61, 173.

lines. (These changes were of the same general kind as those associated with the high temperature transition, but there was no opportunity to follow their progress as the temperature varied.) However, it seems a legitimate deduction that there is another transition at low temperatures in which the structure reverts again to cubic. This would compare with Rochelle salt, which is orthorhombic above 24·5° C. and below − 20° C., and monoclinic in the intervening range [9]. Here the substance is piezo-electric as well as possessing a high dielectric constant between its two "Curie points".

It is interesting to notice the effect of the transition process on individual crystallites.

With one solid solution of (Ba, Sr)TiO$_3$ it was possible to show [10] (1) that the crystallites retained their identity throughout the transition, (2) that the texture just above the transition point was accurately reproducible after intermediate cooling to the tetragonal structure, the texture at room temperature being derived from it and itself closely reproducible, after intermediate heating above the transition point, (3) that a single crystal of the cubic structure inverted to a twin consisting in general of all three related tetragonal orientations, *i.e.* the three orientations in which the tetrad axis coincided with one of the tetrad axes of the original cube. This twinning mechanism is obviously capable of minimising the strains which result from the shrinkage in one direction and extension in another when the tetragonal unit cell is formed from the cube. This type of twinning, which mimics higher symmetry in the crystal as a whole, is a well-known phenomenon in mineralogy. It may be noted that it is very characteristic of the mineral perovskite, where the apparent cubes are actually fine-grained twins, and this is taken as evidence of the existence of a high-temperature form which is truly cubic. Twinning of a similar kind has also been observed in Rochelle salt and KH$_2$PO$_4$ when they invert to a form of lower symmetry on cooling below their dielectric Curie point [11].

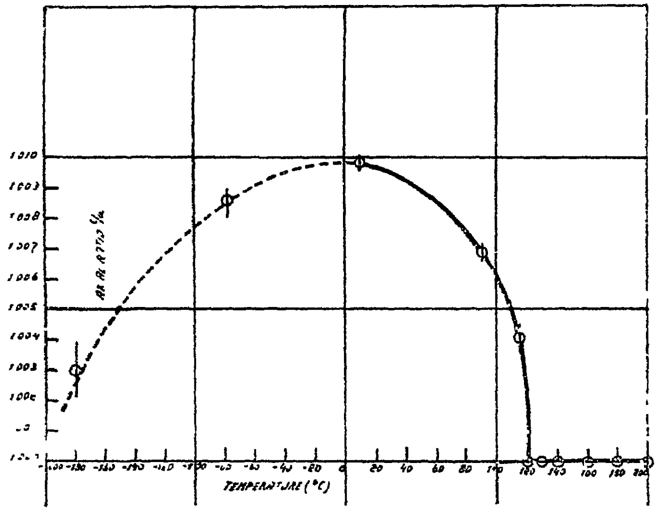


FIG. 5.—Variation of axial ratio of BaTiO$_3$ with temperature.

---

[9] Ubbelohde and Woodward, *Nature*, 1945, 156, 20.

[10] Megaw, *Nature*, 1946, 157, 20.

[11] Ubbelohde and Woodward, *Nature*, 1945, 155, 170.

Summary.

Compounds of the perovskite type having the empirical formula $A^{2+}B^{4+}O_3$ show distorted variants of the ideal perovskite structure when the larger cation A^{2+} is too large or appreciably too small for the geometrical requirements of the structure, the scale being determined by the B^{4+}—O distance. When A^{2+} is too small, a monoclinic distortion of the original cubic lattice occurs (but with an equality of cell edges which implies that it is strictly orthorhombic, though to display this would need a different choice of axes). This change of cell shape is accompanied by a doubling of the cell edge. This is found in $CaTiO_3$, $CaSnO_3$, $CaZrO_3$, and $CdTiO_3$.

$BaTiO_3$ is the unique example with too large an A^{2+} ion. This results in abnormally large Ti—O distances. There is a tetragonal deformation, but no doubling of the cell edge, and, unlike $CaTiO_3$, this cannot be accounted for on purely steric grounds. It is attributed to a set of directed bonds of the Ti atom which are postulated to occur superimposed on the mainly ionic forces of the structure.

$BaTiO_3$ has a λ-point transition about 120° C., above which it becomes cubic. The transition is accomplished with no discontinuous change of spacing, though with a peak or discontinuity in the linear thermal expansion coefficients; it takes place over a range of several degrees.

Below room temperature there is a decrease in the axial ratio which suggests a reversion to the cubic structure at some temperature below —183° C.

Some observations are included on twinning of individual crystallites during the transition, and on crystal recovery near the transition temperature.

I wish to thank Mr. J. A. M. van Moll and the Directors of Philips Lamps Ltd. for permission to publish this paper.

Material Research Laboratory,
Philips Lamps Ltd.,
Mitcham Junction, Surrey.

THE PERMITTIVITY OF POLYCRYSTALS OF THE PEROVSKITE TYPE.

By D. F. Rushman and M. A. Strivens.

Received 1st March, 1946.

The unusually high permittivity of barium titanate (*ca.* 2000) was first reported by Wul[1, 2] who showed that other metatitanates of the perovskite structure, whilst having a lower permittivity than barium titanate, also show the same phenomenon to a certain extent.

Willis Jackson[3] has since shown that no clear conclusions can be reached by consideration of the permittivity at room temperature alone, and that the form of the temperature *vs.* permittivity curve is far more significant. Barium titanate shows a very sharp maximum permittivity at 127° C., but by forming mixed crystals of barium strontium titanate this peak occurs at lower temperatures. Thus the high permittivity of barium

---

[1] Wul and Goldman, *Compt. Rend. Acad. Sci.* (*U.R.S.S.*), 1945, xlvi, no. 4, p. 139.

[2] Wul, *Nature*, 1945, 156, 480.

[3] Willis-Jackson and Reddish, *ibid.*, 1945, 156, 717.

titanate arises partly from the fact that its peak permittivity lies near the usual temperature ranges and Wul's conclusions could only be valid if the permittivities of perovskite structures were compared at equivalent temperatures.

In this paper we have investigated the permittivity *vs.* temperature curves of solid solutions of barium, strontium and lead titanates, and have shown how the form of these is related to the structure of each material.

The structural data used by us have been determined in these laboratories by Dr. Helen Megaw and are in course of publication elsewhere.[4, 5, 6]

Preparation of Samples.

The ideal material for the measurement of permittivity is one having an apparent density equal to its X-ray density. To achieve the lowest possibly porosity very high compacting pressures and high firing temperatures are required, but as the titanates of barium and strontium only begin to sinter appreciably at above 1000° C. and as lead titanate begins to decompose above this temperature, it is very difficult to avoid porous samples. By the addition of small proportions of ceramic fluxes, *e.g.* soapstone or kaolin, it is possible to obtain densities approaching the X-ray value, but it is not known how such additions may affect the dielectric measurements, and we have therefore preferred to work with pure materials and to introduce corrections for the lower apparent density.

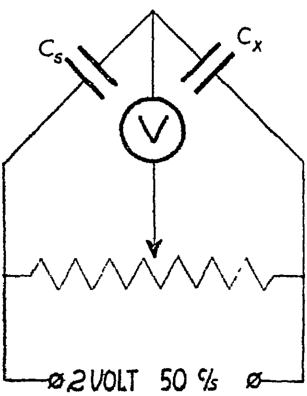


FIG. 1.

The preparation of the samples consisted in milling the required proportions of barium carbonate, strontium carbonate, lead dioxide and titanium dioxide (each of the purest available grade) using a hardened steel mill. After firing the finely divided mixture at *ca.* 900° C. in air to form the titanates, the product was crushed, milled and then pressed in a die at 200 kg./cm.^{2} to give discs about 10 mm. diameter and 1 mm. thick. These discs were then fired at a suitable temperature (1400° C. for mixtures not containing lead; *ca.* 1200° for mixtures with lead titanate) and allowed to cool slowly. After grinding the faces parallel, the discs were measured and then silvered on opposite faces by baking with a silver oxide suspension. In view of their porosity the test condensers thus made were dried at 120° C. and then protected by a lacquer known to have good electrical properties.

Electrical Measurements.—The capacities of the test condensers were determined as a function of temperature using a simple 50 c/s measuring bridge as shown in Fig. 1.

C_x is the unknown capacity, C_s one of a series of fixed condensers of known value, and *V* a valve voltmeter fitted with a "magic eye" as a null point indicator. The test condensers were immersed in a well stirred bath of transformer oil the temperature of which was read by a mercury thermometer to ± 0·1° C. Low temperature measurements were carried out using a pentane/liquid nitrogen bath, the temperature being recorded by a pentane thermometer to ± 2° C.

The errors in the capacity determination amount to ± 2 % except

---

[4] Megaw, *Nature*, 1945, **155**, 484.
[5] Megaw, *ibid.*, 1946, **157**, 20,
[6] Megaw, *Proc. Physic. Soc.*, 1946, **58**, 133.

in cases where the capacity of the sample has dropped to below 100$\mu\mu$ F, in which case stray capacities become serious. Further errors arising in imperfections of the disc and from the measurement of its dimensions bring the total uncertainty in the effective permittivity to ± 2·5 %. Owing to the high permittivity of the samples the edge correction is negligible if the electrodes extend right to the edge of the material.

Porosity Corrections.—Several formulæ have been suggested relating the permittivity of a heterogeneous dielectric to the permittivity of its constituents.[7] In our case the permittivity of the continuous phase is very much greater than that of the discontinuous phase and we proceed as follows from a law derived by Wiener:[8]

$$\frac{\epsilon_A - \epsilon_2}{\epsilon_A + 2\epsilon_2} = x_1\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + 2\epsilon_2}$$

where ϵ_A = apparent permittivity, ϵ_1 = permittivity of dispersed phase (assumed to be spheres), ϵ_2 = permittivity of continuous phase, x_1 = volume fraction of dispersed phase.

In our case $\epsilon_1 = 1$ and $\epsilon_2 \gg \epsilon_1$,

$$\therefore \frac{\epsilon_A - \epsilon_2}{\epsilon_A + 2\epsilon_2} = \frac{-x_1}{2}$$

$$\therefore \epsilon_2 = \frac{2 + x_1}{2(1 - x_1)} \cdot \epsilon_A. \qquad (1)$$

For very low values of x_1 this may be further reduced to:

$$\epsilon_2 = \left(1 + \frac{3}{2}x_1\right)\epsilon_A. \qquad (2)$$

We may evaluate x_1 from the difference between the apparent and the true (X-ray) densities:

i.e.,
$$x_1 = \frac{\rho_X - \rho_A}{\rho_X} \qquad (3)$$

where ρ_A = apparent density, ρ_X = X-ray density.

As in general our porosities have been in the order of 20 % we have taken expressions (3) and (1) for the calculation of the true permittivity of our samples.

Results.

Fig. 2 gives the permittivity of barium titanate as a function of temperature and clearly shows the sharp maximum characteristic of this type of titanate. It was found that the high temperature part of this curve (*i.e.* for temperatures above 127° C.), closely fitted a hyperbolic law of the form

$$\epsilon = \frac{A}{T - T_a} \qquad (4)$$

where A and T_a are empirical constants.

It was also found that strontium titanate over the range − 100° C. to + 100° C. obeyed a similar type of law. As barium titanate at room temperature has a tetragonal modification of the perovskite structure (Megaw,[4] Rooksby[9]) whilst strontium titanate has the ideally cubic modification, it was suspected that the high temperature part of the permittivity *vs.* temperature curve for barium titanate might also be due to a cubic modification.

This was indeed found to be so but a fuller investigation showed that the original supposition that the temperature of the peak permittivity corresponds to the transition temperature of the two modifications was not true. In Fig. 2 is also shown the axial ratio c/a of the $BaTiO_3$ unit cell as a function of temperature. We may distinguish two important temperatures, namely the temperature T_m corresponding to maximum permittivity and the temperature T_c at which c/a becomes unity. For barium titanate these temperatures lie at 127° C. and 119° C. respectively.

---

[7] Lowry, *J. Franklin Inst.*, 1927, **203**, 413.

[8] Wiener, *Leipzig Ber.*, 1910, **62**, 256.

[9] Rooksby, *Nature*, 1945, **155**, 484.

From the above it is to be expected that strontium titanate would show a similar permittivity peak at very low temperatures and that by the gradual replacement of strontium for barium in barium titanate, the value of T_m would be gradually reduced. Again expectations were realised and Fig. 3 shows the effect of increasing $SrTiO_3$ content on the value of T_m for $(Ba/Sr)TiO_3$ polycrystals. If, on the other hand, mixed $(Ba/Pb)TiO_3$ polycrystals are considered, T_m rises with increasing $PbTiO_3$ content, which fact is in accordance with the larger tetragonality ($c/a = 1{\cdot}06$) of lead titanate. Fig. 3 also shows this together with the corresponding cell dimensions for both series of compounds.

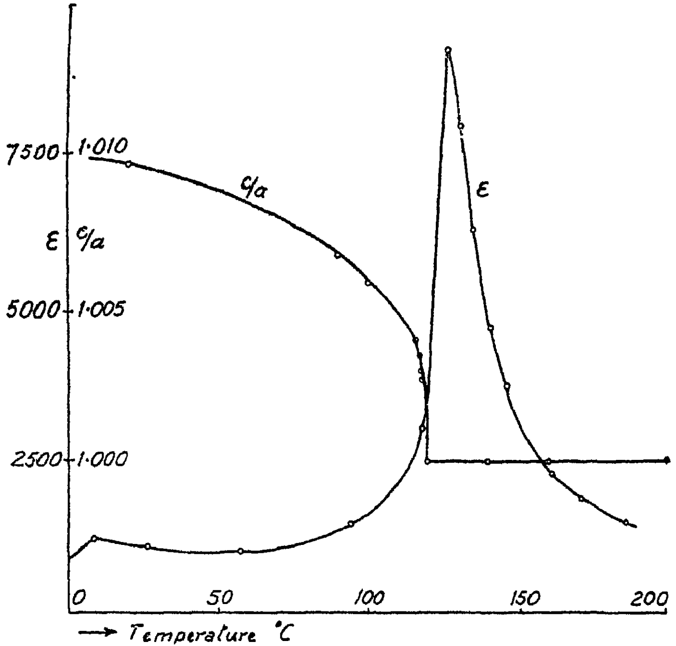


FIG. 2.

The most significant point which arises from the ϵ *vs.* T curves for such mixed crystals is that in each case the high temperature part of the curve will fit eqn. (4) and that although T_s varies with the composition, the values of A are nearly constant. This is strikingly parallel to the decay of paramagnetism above the Curie point of a ferromagnetic material, in which case we should interpret the constancy of A as being due to a constant (magnetic) dipole for different compositions (we neglect here the effect of changing density between the various compositions). Considered in connection with the shape of the axial ratio *vs.* temperature curve, these facts point to a somewhat analogous co-operative phenomenon occurring in these titanates and it is not difficult to seek an explanation.

Discussion.

From a consideration of the crystallographic data it is readily concluded that the titanium ion in barium titanate has a larger volume to occupy than that deduced from the Goldschmidt radii of the titanium and oxygen ions. To adopt a mechanical picture we may say that the titanium ion can " rattle " within its enclosing octahedron of oxygen ions. As the repulsive forces fall off very rapidly with increasing distance it is not difficult to see that the geometrical centre of the oxygen octahedron must represent a local potential energy maximum and that the equilibrium position of the titanium ion must lie to one side of this point. This is illustrated by Fig. 4 which shows the form of the potential energy of a titanium ion as a function of position between two oxygen ions placed at $x = a$ and $x = -a$. In general there will be an infinite number of equilibrium points within the oxygen octahedron which will define a surface enclosing and symmetrical about the geometric centre.

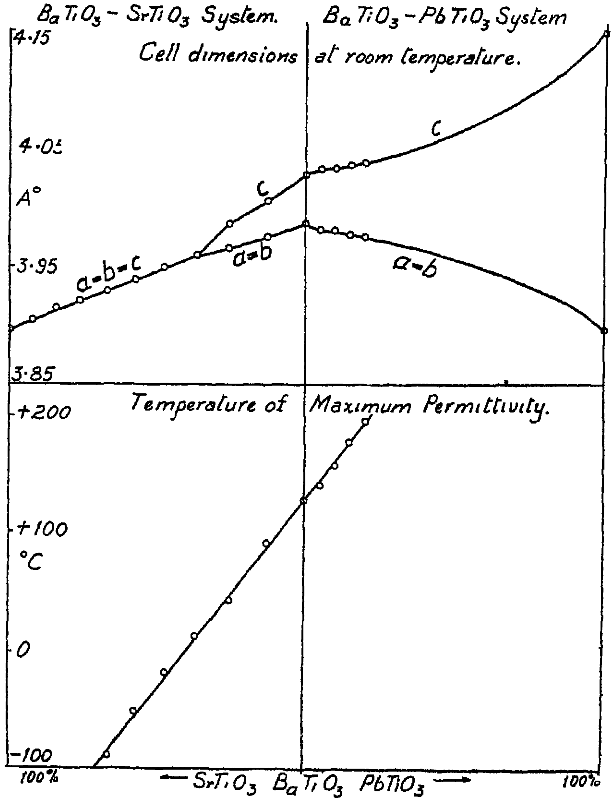


FIG. 3.

As the titanium ion is quadruply charged this displacement from the geometric centre which is also the charge centre of the barium and oxygen ions, will give rise to a dipole $4er_e$ where r_e is the displacement from the centre. We may thus consider barium titanate as an assembly of dipoles of fixed location in a crystal lattice and owing to the strong coupling between neighbouring dipoles via the highly polarisable oxygen ions, it is to be expected that they will be subject to a co-operative rotation-libration equilibrium. The close similarity to the case of Rochelle salt which has H_2O dipoles undergoing libration or rotation is very apparent.

Consider now a perovskite structure of general formula $MTiO_3$ at a temperature above the rotation-libration Curie point. As each dipole is now freely rotating, the mean position of the titanium ion coincides with the geometrical centre of the lattice and from considerations of symmetry it is to be expected that the lattice should be cubic. If now the titanium ion is displaced towards a particular oxygen ion as shown in Fig. 5, *i.e.* if the dipole is librating, there will be a tendency for the oxygens in the plane at right angles to the direction of the displacement to be drawn inwards towards the titanium ion owing to the electrostatic attraction. The structure should then become tetragonal with an axial ratio $c/a > 1$, and as the libration of the dipoles is co-operative it is to be expected that the curve showing c/a *vs.* T should be very similar to the curves of degree of order *vs.* temperature well known in order-disorder phenomena.

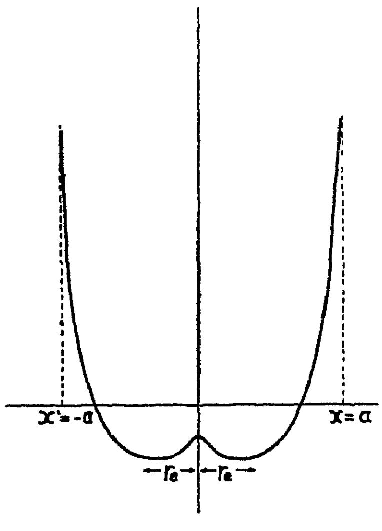


FIG. 4.

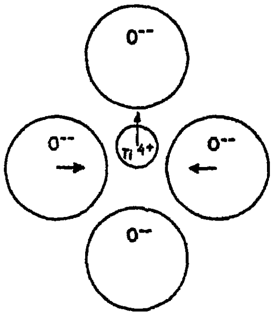


FIG. 5.

Thus the general shape of the c/a *vs.* T curve in Fig. 2 is explained and we may identify T_c with the Curie point for the rotation-libration of the titanium dipoles. Furthermore, as the dipoles are free to rotate above this temperature the usual Debye equation should hold. The application of this is greatly simplified by the fact that the structure is cubic in this temperature range and therefore the Clausius-Mosotti expression for the total polarisability will be valid.

We may therefore write:

$$\frac{\epsilon - 1}{\epsilon + 2}\cdot\frac{M}{\rho} = \frac{4}{3}\pi N\left(\alpha_E + \alpha_A + \frac{\mu^2}{3kT}\right) \quad . \quad . \quad . \quad (5)$$

where μ = value of titanium dipole ($= 4er_e$), α_E = electronic polarisability of unit cell, α_A = atomic polarisability of unit cell, the remaining symbols having their usual significance.

Proceeding now to plot $\frac{\epsilon - 1}{\epsilon + 2}$ *vs.* $1/T$ in the usual way we are able to evaluate μ and $(\alpha_E + \alpha_A)$ for a variety of compositions. The results of this are shown in Table I.

For comparison we have calculated the values of α_E and thus of α_A and α_A/α_E for each composition assuming the polarisability to be additive and taking the values of the refractive index for visible light of the mineral forms of $CaTiO_3$, $BaSiO_3$, $CaSiO_3$ and similar series of compounds as recorded in the International Critical Tables. No extrapolation to the refractive index for light of infinite wavelength has been carried out owing to the paucity of recorded data, and therefore the values of α_E, α_A and α_A/α_E must be regarded as very approximate.

Inspection of the values of μ shows the remarkable fact that the addition of strontium has no marked effect on the amount of displacement of the titanium ion, although the unit cell edge of $SrTiO_3$ is only 3·897 A. as opposed to 3·999 A. for $BaTiO_3$. This decrease of 0·102 A. would be expected to reduce considerably the magnitude of the titanium displacement or even prevent it altogether. The fact that the same dielectric behaviour is observed for pure strontium titanate leads us to believe that even in this compound the titanium is free to " rattle " and we therefore have to attribute to the titanium ion a smaller radius than the usually accepted figure of 0·64 A. This is not such a serious step when we consider that this value has been deduced from structures rather similar to the present case in which the titanium is enclosed in an oxygen octahedron, and that the possibility of " rattling " has been ignored.

TABLE I.

Compositions in Mol %.			μ ($\times 10^{18}$ esu cms.).	r_e (A.).	$\alpha_E + \alpha_A$ ($\times 10^{24}$ cc.).	α_E ($\times 10^{24}$ cc.).	α_A ($\times 10^{24}$ cc.).	α_A/α_E.
$BaTiO_3$.	$SrTiO_3$.	$PbTiO_3$.						
74·5	25·5	—	0·103	0·021	14·9	9·5	5·4	0·57
63·6	36·4	—	0·098	0·020	14·7	9·4	5·3	0·56
52·9	47·1	—	0·108	0·022	14·6	9·3	5·3	0·57
42·8	57·2	—	0·110	0·023	14·4	9·2	5·2	0·57
33·3	66·7	—	0·111	0·023	14·3	9·2	5·1	0·55
24·3	75·7	—	0·098	0·020	14·2	9·1	5·1	0·56
84·3	11·9	3·8	0·148	0·031	15·0	9·7	5·3	0·55
80·8	11·3	7·9	0·145	0·030	15·0	9·7	5·3	0·55
73·1	23·2	3·7	0·146	0·031	14·8	9·6	5·2	0·54
70·2	22·3	7·5	0·144	0·030	14·8	9·6	5·2	0·54
67·2	21·4	11·4	0·139	0·029	14·8	9·7	5·1	0·53
62·4	34·0	3·6	0·139	0·029	14·6	9·5	5·1	0·54

Turning now to consider the region where $T < T_0$ we see that as T is reduced from a high value towards T_0 the permittivity will rise, tending to become infinite when $T = T_0$. At this point the value of $\frac{\epsilon - 1}{\epsilon + 2}$ becomes 1 and the material is infinitely polarisable. This means that a state of permanent polarisation is set up and the material becomes ferroelectric.

The effect of an electric field in this region should be essentially similar to the case of ferromagnetism where domains of spontaneous magnetisation occur. The behaviour of barium titanate and the other mixed titanates in the ferroelectric region is indeed strikingly analogous to the behaviour of ferromagnetic materials. For example, dependence of the permittivity on the applied field and electrostriction are manifested, but a full treatment is made difficult by the dependence of the structure on the temperature and the variations of μ which obviously become significant as the *c* axis grows at the expense of the two *a* axes. This uncertainty in the value of μ in the tetragonal structure, apart from the uncertainty introduced into the Clausius-Mosotti equation by the departure from cubic symmetry, prevents us from calculating the saturation polarisation for such a material. We can only say from experimental observation that breakdown occurs before saturation polarisation is reached.

Summary.

The main features of the dielectric polarisation phenomena occurring in barium titanate and the mixed Ba/Pb/Sr titanates have been presented and discussed. By postulating that the equilibrium position of the titanium ion in these compounds is displaced from the centre of the enclosing octahedron of oxygen ions the main features of the structure *vs.* temperature and the permittivity *vs.* temperature curves have been explained and the existence of ferroelectric properties which are associated with a tetragonal structure is shown.

We wish to thank Mr. van Moll and the Directors of Philips Lamps Limited for permission to publish this paper.

Material Research Laboratory,
Philips Lamps Ltd.,
New Road,
Mitcham Junction, Surrey.

MEASUREMENTS OF THE DIELECTRIC PROPERTIES OF ICE.

BY J. LAMB.

Received 4th March, 1946.

It is well known [1, 2] that ice exhibits a loss maximum in the neighbourhood of 10^4c/s. but little information exists concerning its behaviour at centimetre wavelengths.

The present work was initiated to investigate the possible occurrence of a second absorption maximum in this region of wavelength. The paper describes measurements which have been carried out to determine the dielectric properties of ice under the following conditions.

(1) at a frequency of 10^{10} c/s. over the temperature range 0 to $-40°$ C.
(2) at a temperature of $-5°$ C. over the frequency range 8×10^3 to $1{\cdot}25 \times 10^6$ c/s.

The ice samples used in each experiment were formed from water which had been twice distilled and boiled immediately prior to freezing.

1. Measurements at 10^{10} c/s. over the Temperature Range 0° to $-40°$ C. —A previous paper [3] has described the theory and experimental technique employed in measurements of the dielectric properties of materials at centimetre wavelengths, utilising a cylindrical resonator operating in the H_{01} mode. This method has been used to determine the permittivity and power factor variations of ice at a frequency of 10^{10} c/s. over the temperature range 0 to $-40°$ C.

The variable length H_{01} resonator employed, consisted of a silver plated tube, closed at one end by a tightly fitting piston containing a cavity in which the ice specimen could be formed, and at the other by a non-contact type piston, the axial position of which could be adjusted by means of a sensitive micrometer in the manner shown in Fig. 1.

The resonant response of the system was delineated by variation of its length about the resonant value, and the Q-factor evaluated in terms of the width of the resonance curve at half height, for square law rectification by the detection system.

The permittivity of the dielectric specimen was derived from the

---

[1] Granier, *Compt. Rend.*, 1924, **179**, 1314.

[2] Errera, *J. Physique*, 1924, **5**, 304.

[3] Horner, Taylor, Dunsmuir, Lamb and Willis Jackson, *J.I.E.E.*, 1946, **93**, Part iii, p. 53.

observed change of resonant length which occurred upon its insertion into the system, whilst the loss factor was obtained from the measured Q-factors of the resonator when completely air-filled and when containing the test specimen, respectively. In this particular experiment the length of the tube and the total micrometer travel were such that resonance of the air-filled cavity could be recorded when its length was equal to four or to five half wavelengths.

The resonator was energised, through a small hole drilled in its side, from a reflection type oscillator connected to a length of rectangular section H_{01} waveguide. A similar hole located diametrically opposite this energising one, served to connect the system to the detector, which consisted of a crystal rectifier mounted in a second section of H_{01} rectangular guide. The centre line of the holes was situated approximately at a position of maximum longitudinal magnetic field in the resonator. The inside radius of the tube was 2·37 cm. and for an energising wavelength of 3 cm. this permitted the establishment of the H_{11}, E_{01}, H_{21} and E_{11} modes, as well as the desired H_{01} mode. All these extraneous modes, however, possessed either a radial or longitudinal component of electric field and their excitation to any significant extent was inhibited by means of a pyramid of four resistance wires mounted on a post projecting from the centre of the movable piston.

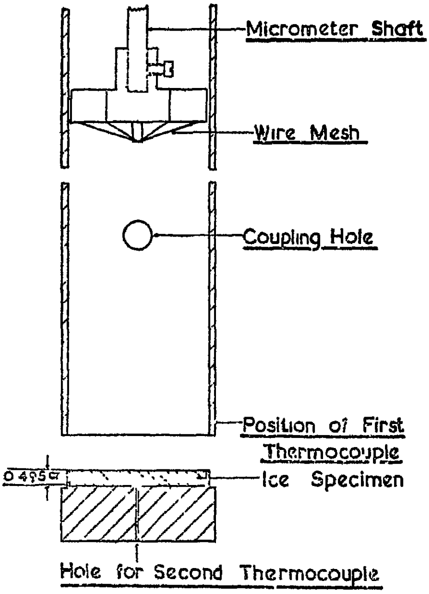


FIG. 1.

The ice specimen consisted of a disc, approximately 0·5 cm. thick, formed in the cavity of the fixed piston shown at the foot of Fig. 1. Its surface was shaped in a plane level with the rim of this cavity, prior to its insertion into the resonator.

The system was cooled by surrounding the lower half of the resonator with powdered solid carbon dioxide which initially lowered the temperature of the ice to about −40° C. The temperature was recorded by means of two thermocouples arranged as in Fig. 1. As the carbon dioxide in contact with the resonator surface evaporated, the system warmed up gradually, but the rate of increase of temperature was sufficiently small to permit measurements to be taken at intervals of a few degrees centigrade. It was not possible, however, to plot a full resonance curve at any one temperature and hence the Q-factors of the system were derived from " spot " readings of the widths of the resonance curves at one half and one fifth of the maximum response, respectively. The Q-factor of the air-filled system was about 24,000 and this decreased to a value between 5,000 and 13,000 on insertion of the specimen, the exact value depending upon the temperature at which the measurements were taken.

In calculating the permittivity value it was necessary to apply a correction to allow for the axial contraction of the tube on cooling, and this was determined from an experiment in which the completely air-filled resonator was cooled under identical conditions. The effective contraction was determined by observing the variations in the resonant setting of the micrometer over the temperature range involved.

It was found that the Q-factor of the air-filled system remained sensibly constant as the temperature was lowered. The effect of any possible hoar-frost formation inside the resonator was therefore not sufficient to be detected.

The points determining the tan δ temperature curve of Fig. 2 were derived from measurements on two separately prepared specimens of ice. The recorded permittivity was $\epsilon' = 3{\cdot}05$ and there was no observable change in this value over the temperature range 0 to $-40°$ C.

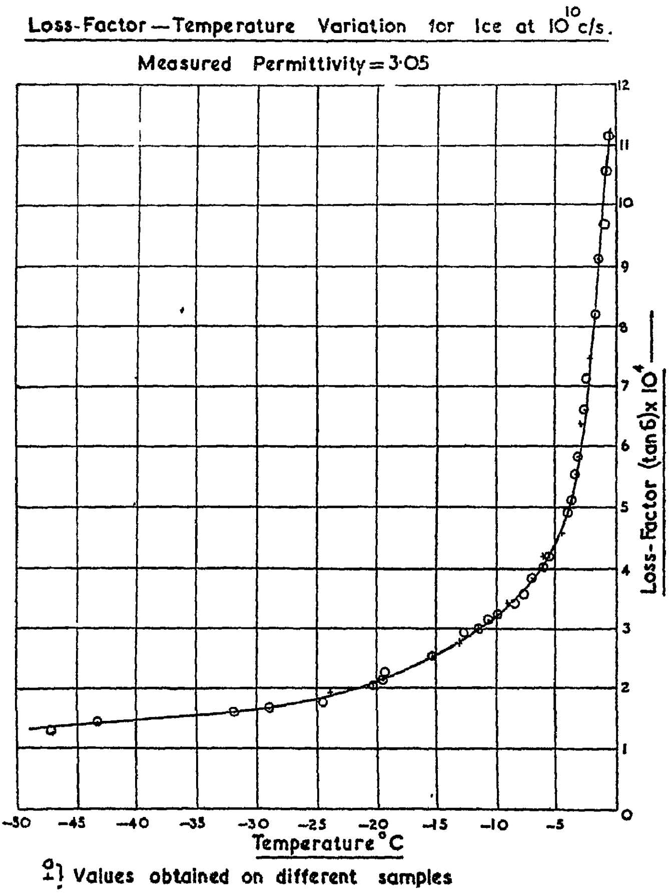


FIG. 2.

It was found from a separate experiment that the rectification property of the detection system conformed to a square law to within 1½ % and it was estimated that the accuracy of permittivity measurement was better than ± 2½ % and that of tan δ evaluation within ± 7 %.

2. The Measuring Technique employed for Frequencies between 8×10^3 and $1{\cdot}25 \times 10^6$ c/s.—The properties of ice over the frequency range 8×10^3 to $1{\cdot}25 \times 10^6$ c/s were measured by utilising a dielectric test set of the form developed by Hartshorn and Ward.[4] The principle of the method is that of capacitance variation in a resonant circuit and a full description of the technique is given in the original paper.

---

[4] Hartshorn and Ward, *J.I.E.E.*, 1936, 79, 594.

The ice specimen to be measured was formed between the circular electrodes of a parallel plate condenser, Fig. 3, and, in the experimental arrangement adopted, this was connected in series with a nominal $50\mu\mu$ F ceramic condenser. The combination was placed across the main condenser of the test set.

The plates of the ice condenser were made of brass and were supported in ebonite blocks, constructed so as to provide the desired separation and alignment of the system. The dimensions were as follows:

Diameter of condenser plates = 3·82 cm.
Plate separation = 0·133 cm.
Theoretical capacitance of this arrangement, neglecting fringing of the electric field at the plate circumference = $7{\cdot}63\mu\mu$ F.

The capacitance of this condenser, when air-filled, was found to be $16{\cdot}5\mu\mu$ F, indicating a total stray capacitance, $C_G = 8{\cdot}9\mu\mu$ F. This consisted of the edge effect due to the capacitance through the ebonite surrounding the brass electrodes (about $4\mu\mu$ F) together with a stray capacitance of about $5\mu\mu$ F to the earth shield which it was necessary to provide round the ice condenser. If C_s is the capacitance of the series ceramic condenser and C_M, $\tan\delta_M$, the measured capacitance and loss factor respectively, of the series combination of this and the ice condenser, then the corresponding values for the latter are given by:

$$\tan\delta_I = \frac{C_s \tan\delta_M}{C_s - C_M(1 + \tan^2\delta_M)} \qquad (1)$$

and

$$C_I = \frac{C_M(1 + \tan^2\delta_M)}{(1 + \tan^2\delta_T)} \cdot \frac{C_s}{C_s - C_M(1 + \tan^2\delta_M)}. \qquad (2)$$

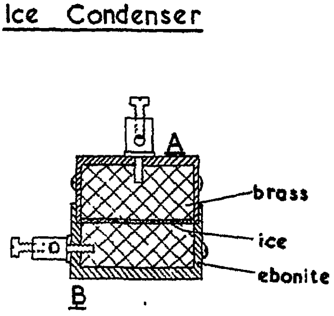


FIG. 3.

The values C_I and $\tan\delta_I$ refer to the ice condenser as a whole and therefore include the effects of the stray capacitance C_G. Hence if ϵ'_1, $\tan\delta_1$, represent the relative permittivity and loss factor of the ice itself, and C_{10} is the theoretical capacitance of the condenser when air-filled ($= 7{\cdot}63\mu\mu$ F) then:

$$\tan\delta_I = \frac{C_I \tan\delta_I}{(C_I - C_G)} \quad . \quad . \quad . \quad . \quad (3)$$

and

$$\epsilon'_1 = \frac{(C_I - C_G)}{C_{10}}. \quad . \quad . \quad . \quad . \quad (4)$$

The series condenser, C_s, was used in order to reduce the equivalent capacitance and loss factor of the measured impedance and for this purpose a value of about $50\mu\mu$ F was found to be convenient. The measured value of C_s was $48{\cdot}6\mu\mu$ F, independent of the frequency, and the corresponding loss factor was of the order 0·001: this was negligible in comparison with that of the ice.

The ice was formed in section (A) of the condenser (Fig. 3) and its surface shaped in a plane level with the surrounding ebonite rim. Section B of the condenser, which was at a temperature just above 0° C., was then placed firmly in position and the whole system cooled to about −10° C. and maintained at that temperature until it was connected to the apparatus. During the experiment the ice condenser was located in a Dewar flask which was closed by a frequentite disc through which the two coupling leads projected. The cooling was provided by a current of air which came from a source of constant pressure and passed through a system of copper coils immersed in liquid air, and then into the Dewar flask. A second current of air at room temperature was also supplied to assist in maintaining a stable temperature.

The temperature was measured by means of a thermocouple which fitted into a small hole drilled in the base of section (B) of the condenser.

The ceramic condenser was connected to the system exterior to the Dewar flask and was thus maintained approximately at room temperature. An earthed shield was placed round the ice condenser to prevent possible power loss in the walls of the flask.

The *Q*-factors of the system, as derived from the widths of a given resonance curve at heights ranging from 0·2 to 0·8 of the maximum, agreed to within ± 3/4 %. The *Q*-factor of the resonant circuit decreased with decreasing frequency and consequently the accuracy of measurement varied over the frequency range. It was estimated that the accuracies of permittivity and tan δ evaluation were each within ± 5 %.

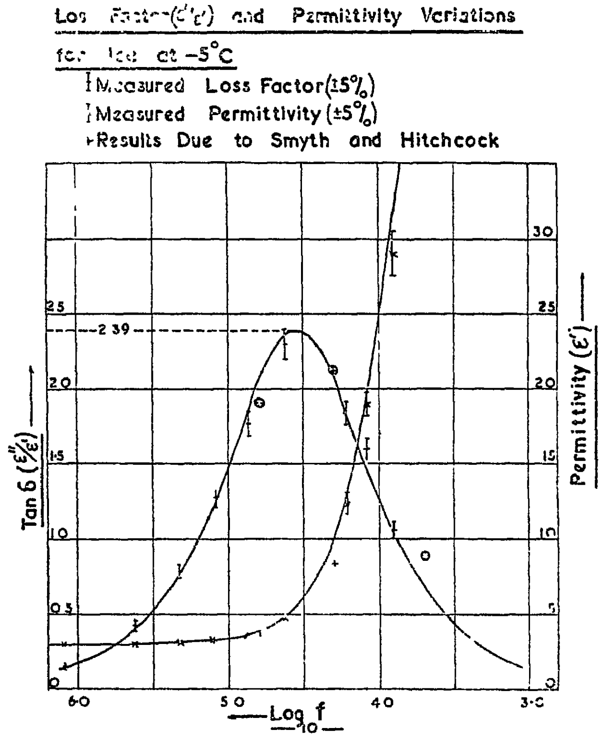


FIG. 4.

The lowest frequency which could be attained with the apparatus was about 8 × 10^3 c/s. and the results of a series of measurements extending from this value to 1·25 × 10^6 c/s. at a temperature of − 5° C. are given in Fig. 4 together with a few values due to Smyth and Hitchcock[5] for the same temperature. By use of the expression (5) and (7) given in section 3, the theoretical curves of Fig. 4 have been drawn as the closest fit to the experimental points.

It can be seen from Fig. 4 that the permittivity of ice at − 5° C. attains a steady value at a frequency of about 4 × 10^5 c/s. The limiting value of the permittivity at high frequencies (denoted by ϵ_∞) obtained from this series of experiments agrees, within experimental error, with the permittivity value, $\epsilon' = \epsilon_\infty = 3{\cdot}05$, derived at 10^{10} c/s. thus indicating that no further absorption peak occurs in the intervening frequency region 1·25 × 10^6 c/s. to 10^{10} c/s.

[5] Smyth and Hitchcock, *J. Amer. Chem. Soc.*, 1932, **54**, 4631.

3. Examination of the Experimental Results.—The modern theory of dielectric behaviour introduced by Debye[6] and modified by Onsager[7] and others, gives expressions for the variations with frequency of the generalised dielectric constant, $\epsilon = \epsilon' - j\epsilon''$, and in the case of a dielectric for which the static dielectric constant ϵ_0 is much greater than the limiting high frequency value ϵ_∞, these take the form

$$\epsilon' = \epsilon_\infty + \frac{(\epsilon_0 - \epsilon_\infty)}{1 + a^2f^2} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

$$\epsilon'' = \frac{(\epsilon_0 - \epsilon_\infty)af}{1 + a^2f^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where f = frequency in c/s. and "a" is a factor directly proportional to the relaxation time of the dielectric. The loss factor of the material is then given by

$$\tan \delta = \epsilon''/\epsilon' = \frac{(\epsilon_0 - \epsilon_\infty)af}{\epsilon_0 + \epsilon_\infty a^2f^2} \quad . \quad . \quad . \quad (7)$$

and hence when this is plotted against $\log f$ as abscissa, the resulting curve is in the form of an inverse coshine relation. The loss factor attains a maximum value for a frequency

$$f_m = \frac{1}{a}\sqrt{\epsilon_0/\epsilon_\infty}. \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The corresponding permittivity value is

$$\epsilon' = \frac{2\epsilon_0\epsilon_\infty}{\epsilon_0 + \epsilon_\infty} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

and the maximum loss factor,

$$\tan \delta_{max} = \frac{1}{2}\frac{(\epsilon_0 - \epsilon_\infty)}{\sqrt{\epsilon_0\epsilon_\infty}} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Strictly this analysis is only applicable to gases and dilute solutions where the interactions between molecules can be neglected, but in the case of ice the experimentally derived values for the loss factor are found to lie on the theoretical curve given by eqn. (7) within the accuracy of measurement. This theoretical curve (Fig. 4) has been obtained by assuming the value of the static dielectric constant, $\epsilon_0 = 74{\cdot}6$, given by Smyth and Hitchcock,[5] and by taking the experimentally determined value $\epsilon_\infty = 3{\cdot}0$ for the limiting high frequency permittivity. The maximum loss factor occurs at a frequency of $3{\cdot}5 \times 10^4$ c/s. and the corresponding value of "a" is $1{\cdot}4 \times 10^{-4}$ secs. The theoretical permittivity curve obtained from eqn. (5) is also shown in Fig. 4. The measured permittivity values at the lower frequencies are slightly displaced from this curve. This may be due to the poor accuracy of measurement at the lower frequencies and/or to a possible error in the value of the static dielectric constant assumed in the prediction of the theoretical curve.

If the loss-factor at 10^{10} c/s. were due only to the component resulting from the absorption with its maximum value at $3{\cdot}5 \times 10^4$ c/s. then for a temperature of $-5°$ C. it would be given as:

$$\tan \delta = \frac{(\epsilon_0 - \epsilon_\infty)af}{\epsilon_0 + \epsilon_\infty a^2f^2}$$

$$\bumpeq \frac{(\epsilon_0 - \epsilon_\infty)}{\epsilon_\infty af} \quad \text{since} \quad \epsilon_\infty a^2f^2 \gg \epsilon_0 \quad \text{for} \quad f = 10^{10}\ \text{c/s}$$

$$= 0{\cdot}17 \times 10^{-4} \quad \text{since} \quad a = 1{\cdot}4 \times 10^{-4}\ \text{secs.}$$

The measured loss-factor under these conditions was, in fact, $0{\cdot}45 \times 10^{-3}$, about 27 times larger than the above value. This discrepancy cannot be accounted for by any error in the measurements and, in the absence of further experimental evidence at still higher frequencies, no satisfactory explanation has yet been established. It may possibly be associated with a second absorption maximum occurring at some frequency greater than 10^{10} c/s. or with the presence of small amounts of impurity in the ice samples.

---

[6] Debye, *Polar Molecules*, Chem. Catalog. Co.

[7] Onsager, *J. Amer. Chem. Soc.*, 1936, 58, 1486

The work described formed part of a programme of extra-mural research carried out on behalf of the Ministry of Supply. The author's thanks are due to Professor Willis Jackson and also to the Chief Scientific Officer of the Ministry for permission to publish the paper.

Electrotechnics Department,
Victoria University, Manchester.

GENERAL DISCUSSION

Messrs. W. Reddish, W. Plessner and **Prof. Willis Jackson** (*Manchester*) (*communicated*) : The results of the experimental work which we have carried out during recent years, in co-operation with the British Pottery Research Association and with W. H. Taylor, on the barium-strontium titanate series serve generally to confirm those described by Rushman and Strivens. There are, however, a number of points of difference which it is appropriate to report here pending a fuller publication of our own work.

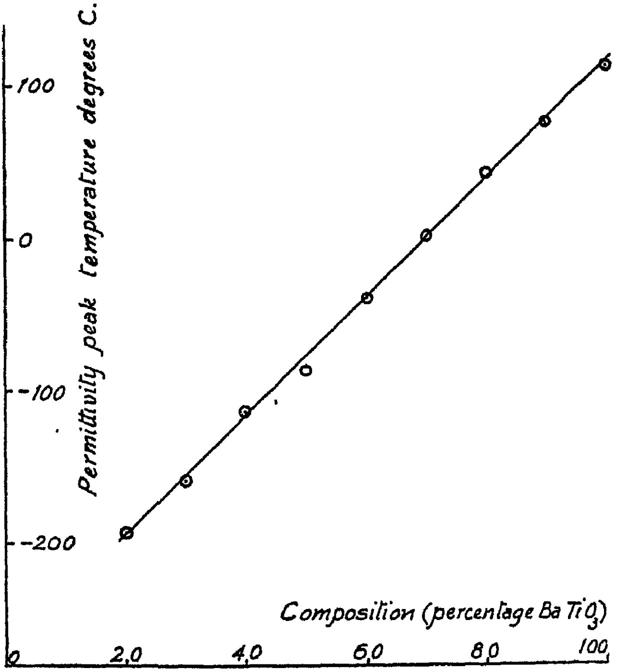


FIG. (*a*).—Permittivity peak temperature-composition for barium strontium titanate solid solutions.

In the first place there is a considerable discrepancy between the variation of the peak permittivity temperature with composition as observed by ourselves—see Fig. (*a*) above—and that shown in Rushman and Strivens' Fig. 3. In both cases the variation is substantially linear but for corresponding specimens our peak permittivity temperatures become progressively lower than theirs with increasing strontium content.

Secondly in seeking to derive the true permittivity, associated with the various compositions, from the observed permittivity values by allowing for the effect of porosity, we have found the logarithmic mixture formula due to Lichtenecker and Rother to be the most satisfactory. This formula states that the permittivity ϵ of a mixture of components of individual permittivities $\epsilon_1, \epsilon_2, \ldots \epsilon_n$ and volume fractions $v_1, v_2, \ldots v_n$ is given by

$$\log_{10}\epsilon = \sum_n v_n \cdot \log_{10}\epsilon_n.$$

The result of applying this formula to a series of specimens of the same basic composition but of varying porosity x is shown in curve (1) Fig. (*b*). It leads to a value of 12,500 for the true permittivity of the sample. Application of Rushman and Strivens' equation (2) to our highest permittivity specimen leads to the straight line (2) which shows no significant correlation with the remainder of our results.

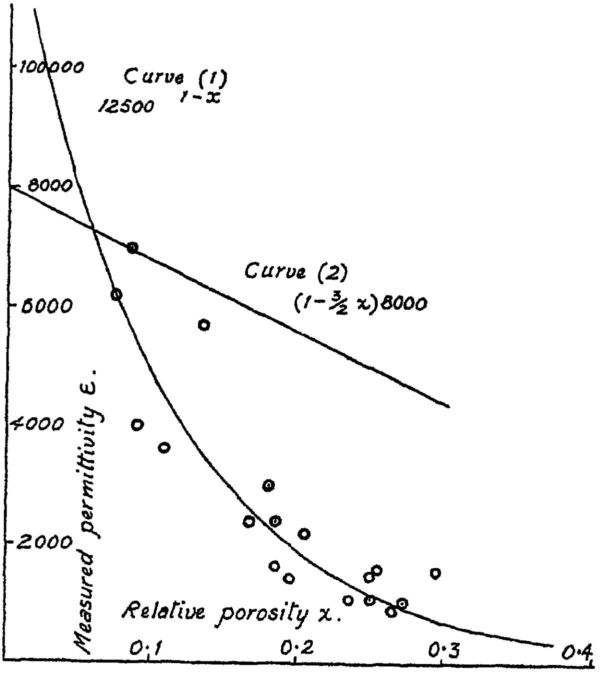


FIG. (*b*).—Relation between measured permittivity and porosity for 72 : 28 (Ba : Sr) TiO_3 samples.

The results shown in Fig. (*c*) are of interest in connection with the statement by Rushman and Strivens that the permittivity of barium-strontium titanate ceramics is affected by the strength of the applied electric field. This figure shows the effect of a superposed D.C. field on the permittivity of a 70 : 30 barium-strontium titanate specimen as measured with weak A.C. field, at a frequency of about 1 Mc./sec.

The physical explanation presented by Rushman and Strivens is of considerable interest, though their dipole conception is difficult to justify notwithstanding the support which the data of their Table I might appear to afford it. It would be helpful if they could offer an explanation of the "tableland" of permittivity which exists on the low temperature side of the permittivity peak and of the fact that its temperature range shortens in the barium-strontium titanate series with increasing strontium content.

Dr. Helen Megaw (*London*) (*communicated*): Dr. Rushman and Mr. Strivens suggest that in the transition from cubic to tetragonal the Ti ion changes from rotation to libration, from moving on a complete sphere to moving on a spherical cap of the same radius and symmetrical about the tetrad axis. This should be accompanied by a *decrease* of volume, due to a decrease in the effective volume of the Ti ion; the *c*-dimension would remain constant, while the *a*-dimension decreased. Experimentally there is *no change* of volume at the transition point, and *c* increases while *a* decreases.

The suggestion that the Goldschmidt radius of the Ti ion is seriously too high is quite untenable on crystallographic grounds. It is arguable that the Goldschmidt radii (derived empirically from crystal geometry) are not of significance for the electrical treatment; but if so, the electrical theory cannot be used to "correct" the radii of individual ions that do not happen to fit it.

Professor E. Bauer (*Paris*) said: I think it is better not to speak of dipoles in the phenomena which Rushman and Strivens have observed; they do not seem due to the rotation of an electric moment, but to the translation of a moving ion from one equilibrium position to another in a lacunar crystal as Miss Megaw has described.

Professor E. Bauer (*Paris*) (*communicated*): When you consider such ion translations, you have no longer to distinguish between Onsager cavity- and reaction-field. Both fields tend to increase the distance between positive and negative ions. In this case we may foresee a "$\frac{4}{3}\pi$ catastrophe," which might well explain the dielectric properties of Ba titanate.

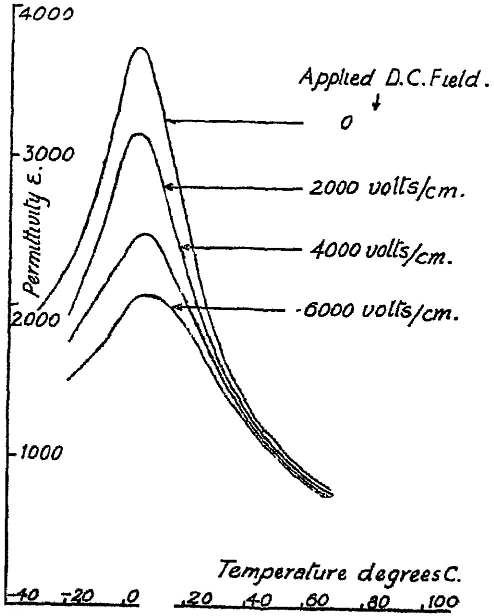


FIG. (*c*)—Effect of superimposed D.C. field on 1 Mc/sec. permittivity—70:30 (Ba:Sr) TiO_3.

Are the "ferro-electric" properties of the Rochelle-salt and KH_2PO_4 due to dipole rotation or to ion displacement? The mechanism of the spontaneous polarisation is still very much discussed. I, personally, believe that it *can* be a dipole rotation. It seems rather difficult, in the case of KH_2PO_4, where the dipolar ion $H_2PO'_4$ is not dissociated, to imagine something different from a dipole rotation.

Dr. W. H. Taylor (*Cambridge*) said: The papers by Miss Megaw and Messrs. Rushman and Strivens deal with the perovskites. The materials examined are (potentially, at least) of great technical and commercial importance, but the attempt to explain their peculiar behaviour is of the greatest interest from the purely scientific point of view. In Rochelle salt and KH_2PO_4 it is supposed that the ferroelectric properties are a consequence of the nature of the H—bonds, which produce dipole effects; but the H atom cannot be located directly by X-ray methods. In the perovskites, by contrast, the ferroelectric behaviour cannot be attributed to H—bond effects, and it may therefore be possible to determine by direct X-ray analysis the precise nature of the structural feature with which ferroelectric behaviour is associated.

We are in full agreement with the conclusions reported in Miss Megaw's paper. Thus we also have established the fundamental distinction between those perovskite-type structures which show a departure from the ideal cubic symmetry because the central ion is too small (*e.g.* $CaTiO_3$), and those which show the type of distortion for which $BaTiO_3$ may serve as an example. We have also applied " single " crystal methods to complex aggregate fragments of the mineral perovskite itself ($CaTiO_3$),and have shown the structure to be orthorhombic. We have observed, as has Miss Megaw, that anomalous intensity-changes in X-ray reflections are associated with the transition from tetragonal to cubic symmetry in the Ba, Sr titanates, and by direct microscopic observations we have been able to demonstrate some of the geometrical relationships between the " subcrystallites " formed by breakdown of minute single crystals when the symmetry is reduced from cubic to tetragonal.

Messrs. Rushman and Strivens have attempted to construct a working hypothesis which could account for the observed electrical properties in terms of the structural data provided by Miss Megaw's work. Mr. Kay and I feel that more accurate knowledge of the structural detail is required to test the validity of this hypothesis, and we are continuing our own work with this aim in view. In the meantime, it is perhaps sufficient to suggest that the experimental evidence does not appear to warrant the introduction of an effective radius for Ti^{+4} different from the usually accepted " Goldschmidt " value. We may also point out that there appears to be an internal inconsistency when Messrs. Rushman and Strivens derive a dipole moment characteristic of the displaced Ti ion; for the argument is based on the fact (Rushman and Strivens' Table I) that in the Ba—Sr titanates the moment is independent of the proportion of Sr present, whereas the same Table also shows that a very small proportion of $PbTiO_3$ changes the Ti-moment by 50 %. I believe that the development of any really detailed theory must await more complete and more accurate structural data, which Mr. Kay hopes to provide.

Mr. D. M. Ritson (*Oxford*) (*communicated*): I should like to draw attention to the theory formulated by Fowler [1] to explain the phenomena associated with Rochelle salt. The theory was remarkably successful in explaining the phenomena that had been found experimentally The formula put forward by Fowler was as follows :—

$$\epsilon = 1 + \frac{4\pi}{\gamma} \cdot \frac{T^{\circ}_0 f.(T_L/T)}{T - T^{\circ}_0 f.(T_L/T)}.$$

It contained two critical temperatures T°_0 which was the temperature for the onset of a co-operative state of self-polarisation and T_L which in Fowler's theory was the critical temperature for the onset of free rotation. (It is obvious that rotation need not be involved, but some other phenomenon such as the onset of long distance order would be mathematically analogous.)

One would tentatively, in the case of barium titanate, identify the temperature of 127° C. with T°_0 and the temperature of 119° C. with T_L. Between 119° C. and 127° C., one would expect barium titanate to be ferroelectric. Below that temperature one would expect the material to be a normal dielectric. This seems to be experimentally confirmed in the work of Wul.[1] He found at room temperature that no frequency or hysteresis effects appeared to exist.

---

[1] Fowler, *Statistical Mechanics*, p. 817-823.

It further seems probable that the crystal is anisotropic, and that along a different axis the dielectric constant again peaks at a low temperature. This would explain the occurrence of a subsidiary maximum.

Dr. G. H. Jonker (*Eindhoven*) said : 1. At first sight it seems very strange that the high temperature curve of ϵ obeys two laws at the same time :

$$\epsilon = \frac{A}{T - T_C} \quad \text{and} \quad \frac{\epsilon - 1}{\epsilon + 2} = \frac{4}{3}\pi N_1\left(\alpha_E + \alpha_A + \frac{\mu^2}{3kT}\right).$$

The former of these laws seems to be the Curie-Weiss law of magnetics and if A has the same significance as in this law we calculate from $A = \frac{N\mu^2}{3k}$ the very high value of $\mu = 15{\cdot}10^{-18}$, whereas Rushman and Strivens calculate from the Clausius-Mosotti formula $\mu = 0,\ 1{\cdot}10^{-18}$. The latter value must be right as it is in conformation with crystallographic data.

Also with the adjacent table it can easily be shown that the high value is wrong

As all the titanates have very high electronic and atomic polarisabilities we get in crystals without dipoles

$$\epsilon \sim 100 \quad \text{or} \quad \frac{\epsilon - 1}{\epsilon + 2} \sim 0{\cdot}97.$$

Only a very small dipole contribution is needed to get the high value of $\epsilon = 1000$.

ϵ.	$\frac{\epsilon - 1}{\epsilon + 2}$.
8	0·70
18	0·85
48	0·94
98	0·97
500	0·994
1000	0·997

The formula $\epsilon = \frac{A}{T - T_C}$ has another meaning as it can be derived from the Clausius-Mosotti formula in the form

$$\epsilon - \epsilon_0 = \frac{4\pi N_1\mu^2}{3k(1 - \frac{4}{3}\pi N_1\alpha)^2} \cdot \frac{1}{T - \dfrac{4\pi N_1\mu^2}{9k(1 - \frac{4}{3}\pi N_1\alpha)}},$$

where $\alpha = \alpha_E + \alpha_A$ and $\epsilon_0 =$ the dielectric constant of crystals without dipoles. As ϵ is very high ϵ_0 can be neglected.

2. If a choice between the Clausius-Mosotti and the Onasger formula has to be made, the former should be taken, as the latter leads to an abnormally high value of the dipole moment. Of course some improvements can be made. As the mean orientation of the dipoles is the direction of the applied electric field, the dipoles are coupled to this direction by that part of the reaction field that is caused by the surrounding dipoles (see the paper of Bauer and Massignon) and this part of the reaction field must be taken in account for the local field. Obviously, in the case of $BaTiO_3$, the total reaction field must not be calculated with the macroscopic ϵ, as the dipole is in the first place surrounded by a layer of O ions and at greater distance by layers of Ba ions and Ti dipoles. There is a possibility here of making a calculation with Kirkwood's theory.

3. Another manner to estimate μ is with an approximate T_C relation of Heisenberg : $kT_C \approx \frac{z}{2}A$, where $z =$ the co-ordination number of the ion for its Ti neighbours and $A =$ an average interaction energy between two neighbouring Ti dipoles. A rough estimation of this energy leads indeed to values of μ of the same order as can be expected from crystallographic data.

Dr. J. H. van Santen (*Eindhoven*) said : Dr. Rushman is using a law derived by Wiener for the permittivity of a dielectric in which spherical holes occur.

---

[1] Wul and Goldman, *Compt. Rend. Acad. Sci.* (*U.R.S.S.*), 1945, 46 (4), 139.

With Mr. Polder I have investigated the more general case of a dielectric with permittivity ϵ_e in which ellipsoidal parts of permittivity ϵ_i occur. We consider here only the case where these parts are vacuum or air-filled holes with $\epsilon_i = 1$. It was not possible to obtain a general exact expression for the macroscopic dielectric constant ϵ. Only in the case of a small volume fraction of interstices exact formulæ could be obtained.

It was found that the shape of the ellipsoidal holes had an influence which increased with ϵ. In the case of a very small volume fraction x of holes we obtained, assuming a random orientation :

$$\text{spherical holes } \epsilon = \epsilon_e \left(1 - \tfrac{3}{2}x\right)$$
$$\text{flat disc-like holes } \epsilon = \epsilon_e \left(1 - \tfrac{1}{3}\epsilon_e x\right)$$
$$\text{long needle-like holes } \epsilon = \epsilon_e \left(1 - \tfrac{5}{3}x\right)$$

It could be proved that ellipsoidal holes had the minimum effect in diminishing ϵ, when the ellipsoids became spheres. From this limiting case we may conclude that a small deviation of the spherical shape will not have much effect.

If a sample has been obtained by compressing a powder, we might assume that the interstices are wedge- or disc-like because the crystals are limited by their natural cleaving faces. So we expect a low value of ϵ. If however the sample has been prepared by a sintering process, we might expect by some surface tension argument that the holes obtain a more or less spherical form.

It might be possible to explain in this way a part of the so-called sintering effect, *i.e.* the fact that for equal volume fraction of holes the permittivity (and in the magnetic case also the permeability) of a sample obtained by compression of a powder is much lower than the permittivity of a sample prepared by a sintering process. This effect is not important in the case of low ϵ materials, but in the case of these titanium compounds it is important.

Mr. M. A. Strivens (*Mitcham*) (*communicated*): Professor Bauer queried our use of the term " dipole " and asked whether the effect was not one of ionic polarisation.

We have used the term *rotating titanium dipole* in our simplified statement of the theory, because we believe it embodies the essential ideas which we wish to put forward. We realised from the beginning and have stated in our paper that the titanium ion has an abnormal freedom of movement and this may be of two kinds.

If the titanium ion occupies the geometric centre of the unit cell, *i.e.* if the latter represents a potential minimum, the vibrations of the titanium ion will constitute a normal atom polarisation. If the titanium ion is permanently displaced from the geometric centre, that is if the latter represents a potential hill, it cannot execute normal vibrations, unless it has simultaneously sufficient thermal energy to pass through the potential barrier.

Instead it may traverse a more devious path between two or more characteristic positions. Since the charge centre of the oxygen and barium ions coincide with the geometric centre, this movement of the titanium ion is mathematically analogous to the rotation or libration of a permanent dipole. However since this mechanism is physically distinct from the rotation of molecule-like nitrobenzene it would perhaps be preferable to use a term such as " ionic dipole " for this new effect.

Dr. Rushman (*Mitcham*), in reply to Mr. Reddish, said : If we evaluate T_C for strontium titanate by extrapolation of the Clausius-Mosotti-Debye formula to $\epsilon = \infty$ we obtain a value of 12° K. which is in reasonable agreement with Mr. Reddish's value for T_m. The difference between our extrapolated value of T_m and his may possibly be due to a difference between the frequency at which the measurements were carried out 50 c./s. in our case and 1·5 mc./s. (?) in his. If this is so it would appear that T_m is somewhat frequency dependent for such compounds.

Mr. Reddish asked how the rapid lowering of the temperature corresponding to the permittivity peak by the addition of strontium titanate to barium titanate could be correlated with the increase in this temperature by the addition of lead titanate.

We regard the Curie temperature T_C, *i.e.* the temperature at which $\epsilon \rightarrow \infty$ and c/a become 1, as of greater fundamental significance than the temperature of maximum permittivity T_m. In general these two temperatures lie close together and arguments concerning the general trend of T_C with changing composition will be equally valid for T_m.

As $\frac{\epsilon - 1}{\epsilon + 2} \rightarrow 1$ as $T \rightarrow T_0$ we may derive from the Clausius-Mosotti-Debye equation :—

$$T_C \quad \frac{4\pi N\mu^2\rho}{3k\{3M - 4\pi N\rho(\alpha_A + \alpha_E)\}},$$

from which is seen that T_C is strongy dependent on $(\alpha_A + \alpha_E)$ or alternatively on the molecular refractivity of the various compounds. The molecular refractivities concerned are :—

$PbTiO_3$—28·9
$BaTiO_3$—24·6
$SrTiO_3$—22·5

The changes of T_c with a composition are thus in accordance with the proposed explanation of the phenomena.

Dr. Megaw queried the validity of our conclusion as the Goldschmidt tolerance factor was less than one in the case of $SrTiO_3$ and therefore "rattling" of the titanium ion could not take place.

The concept of Goldschmidt radii is an idealised one which is eminently suited to the static spatial relationships in crystal structure but it would be very surprising if such a simple concept were valid when considering the dynamic processes amongst the ions themselves. Our conclusions in no way vitiate the Goldschmidt radius as a pointer to the *crystal structure* and dimensions and we feel that the concept of a "rattling" titanium ion is not inconsistent with a tolerance factor less than one. That the tolerance factor is not to be taken too rigidly is evident from Dr. Megaw's paper which quotes 0·98 as the *t*-factor for $PbTiO_3$, whereas this is more tetragonal than $BaTiO_3$ with the higher *t*-factor of 1·02.

Dr. F. C. Frank (*London*) (*communicated*): It was surprising to hear no reference in the discussion to the theoretical work on the electrical cooperative effects, originating with Fowler.[3] Though he wrote of dipoles, there appears little doubt that his work is general enough to account in principle for practically all of the phenomena described to us. As this is a case in which the polarisation is largely electronic, and for the rest depends on very small displacements of ions, it is one in which the Mosotti law can be taken as basic with more than usual confidence. Indeed we may perhaps regard the Curie point of barium titanate as the most genuine example of a "$4\pi/3$ catastrophe". On broad grounds of theory and analogy, it is to be anticipated that progressive substitution of strontium, which lowers the upper Curie point, will simultaneously raise the lower one, which may be useful for bringing it into a more convenient range for study.

When I wrote on the subject in 1936,[4] I indicated, *inter alia*, that to secure a high dielectric constant one should choose substances having the right ratio of ionic radii for close packing, and containing ions with a high ratio of polarisability to volume, so as to secure a high value of

$$p_e = (n^2 - 1)/(n^2 + 2).$$

This however must not be overdone, or electronic dissociation becomes too easy and the substance is prone to behave as an electronic semiconductor:

---

[3] Fowler, *Proc. Roy. Soc., A*, 1935, **149**, 1 and *Statistical Mechanics. See also* Mueller, *Ann. New York Acad. Sci.*, 1940, **40**, 321.

[4] Frank, *Trans. Faraday Soc.*, 1937, **33**, 513.

$p_e = 0{\cdot}7$ is about as high as one dare go, and already dangerous. It is then necessary to have a high atomic polarisation to make $(p_e + p_a)$ up to something near to 1. To secure this one needs highly charged cations contiguous with "soft" anions. Ti^{4+} and O^{2-} satisfy the requirements eminently, but at that time I would have regarded Ba^{2+} as a relatively unprofitable constituent: I failed to foresee that by expanding the structure it could reduce the constraint on the titanium ions. This is a most satisfactory effect, because by increasing p_a and reducing p_e, it secures the desired high total polarisation with less danger of semi-conduction.

In the foregoing remark, p_a has been interpreted as polarisation arising from atomic (or ionic) displacement. This interpretation must be distinguished from that based on the usual $(p_e + p_a + p_0)$ break-up of the total polarisation according to dependence on temperature and frequency: p_a, according to the former definition, is only independent of temperature when the vibrations are harmonic. The point is well illustrated by considering displacement polarisabilities α_a for a number of differently shaped one-dimensional potential energy wells, with progressive change from V-shaped, through parabolic and rectangular wells to those with a double minimum, the extreme being a two-position model. In the following, the ionic charge is Ze, x is a linear co-ordinate, d a characteristic distance and V_0 a characteristic potential energy: α_a is calculated classically for a field parallel to the x-axis. The results are directly applicable to wells of orthorhombic symmetry defined by multiplying together similar potential functions of x, y, and z.

V-shaped well :

Potential energy : $V = V_0|x|/d$

$$\alpha_a = 2kT(Zed)^2/V_0^2$$

Parabolic well :

$$V = V_0(x/d)^2$$
$$\alpha_a = (Zed)^2/2V_0$$

Rectangular well :

$$-d < x < d,\ V = 0 : |x| > d,\ V \gg kt$$
$$\alpha_a = (Zed)^2/3kT$$

Two position well :

$$x = \pm d,\ V = 0 : |x| \gtrless d,\ V \gg kT$$
$$\alpha_a = (Zed)^2/kT$$

Double parabolic well :

$$x > 0,\ V = V_0(x-d)^2/d^2 : x < 0,\ V = V_0(x+d)^2/d^2$$
$$\alpha_a = (Zed)^2\left(\frac{1}{kT} + \frac{1}{2V_0} + \frac{1}{\sqrt{\pi V_0 kT}} \cdot \frac{\exp.\,(-V_0/kT)}{1 + \text{erf.}\sqrt{V_0/kT}}\right)$$
$$\approx (Zed)^2(1/kT + 1/2V_0)$$

It will be noted that in the first model the polarisability increases with rise of temperature: in the second it is independent: and in the remaining three it diminishes as temperature increases. Any one of the latter may be taken as an illustrative model for the titanium ion in barium titanate.

Since the displacement of an ion is dimensionally equivalent to a dipole, it seems at first sight that Rushman's reference to dipoles, where none exist, is purely a matter of nomenclature, unwise, but harmless. Further consideration shows this is not the case. This is best illustrated in terms of a model and assumptions for the barium titanate case, which have been discussed at the meeting. If the titanium ion is enclosed in an approximately cubical box, the faces of which are the surfaces of six oxygen ions, and if a co-operative field along the tetrad axis holds it in contact with the upper face of the box, there may be no reduction of freedom to move on the other two axes except that due to reduction in the *a*-spacing. This is quite different from a dipole oriented by a powerful field which has greatly reduced polarisability for perpendicular directions as well. This, of course, is not in disagreement with the crystallographic statement that there can only be a displacement parallel to the tetrad axis : for the crystallographic analysis does not reveal displacements which are not repeated uniformly from cell to cell, save as a small and symmetrical change in atomic form factor. Thus the major polarisability between the Curie points, when there is spontaneous polarisation along the tetrad axis, may be transverse to this axis, and in these directions less than catastrophically co-operative—hence linear and non-retarded. If other related systems can be found, it may be possible to find further Curie points associated with the loss of further degrees of randomness.

Dr. Helen Megaw (*London*), in reply, said : Barium titanium oxide is of particular interest for the theory of dielectrics as a structure containing no obvious dipoles. It is isodesmic ; that is to say, the forces throughout are of the same kind and same order of magnitude, and there is no segregation into separate molecules or molecule-ions. For this reason it is correct to call it barium titanium oxide rather than barium titanate (though the latter is convenient and not seriously misleading) ; it is a double oxide and not the salt of a complex ion.

With regard to Dr. Jonker's suggestion, I had considered and rejected the Zachariasen radii. While they give an excellent fit with observed interionic distances for binary compounds, for ternary compounds they show much worse deviations than the empirical Goldschmidt set. They involve certain theoretical assumptions about the lattice forces which may not always represent justifiable approximations. The apparently anomalous relation between *t*-value and structure in the lead compounds is more likely to be explained by the partly homopolar character of the Pb—O bonds. I should not be surprised if the occurrence of polymorphous forms explained the discrepancies between the observations of different workers on the structure of these compounds.

I wish to question the emphasis on rotation in connection with λ-points. There is no rotatable group which can be involved in the α-β transition in quartz, or in the ferro-electric KH_2PO_4, and it is not now taken as certain that rotation occurs in other cases where it was once assumed.

While it is formally possible to postulate a rotating dipole in $BaTiO_3$, it does not seem physically very plausible. I do not think that any description depending on vibrations or rotations of the Ti ion can be adequate without some knowledge of the wave-functions of the ion. I believe that such a wave-mechanical treatment is needed to explain the high dielectric constant and high refractive index that are so characteristic of titanium compounds ; and that an explanation of the abnormal properties of $BaTiO_3$ would then follow from the abnormal volume available to the Ti ion.

Finally I should like to express my thanks to Dr. Rushman for introducing me to this problem and for his continued friendly co-operation throughout the work, in spite of our difference of approach to its theoretical interpretation.

Professor A. von Hippel and **Dr. R. G. Brekenridge** (*Massachusetts*) (*communicated*) : The electrical studies of Rushman and Strivens and the crystal structure investigations of Megaw on the perovskite type of crystals offer an interesting complement to work carried out at the Laboratory for Insulation Research of the Massachusetts Institute of Technology during the war for the O.S.R.D. under contract OEMsr-191. The results have been published in two reports : *High Dielectric Constant Ceramics* NDRC, Div. 14, No. 300, August 1944) and *Titania Ceramics II* (NDRC, Div. 14, No. 540, October 1945). Both reports can now be obtained without restriction from the Publication Board of the U.S. Department of Commerce. A paper summarising the major part of these investigations is scheduled for the Nov. issue of *Ind. Eng. Chem.* and was presented at the April 10th meeting of the American Chemical Society in Atlantic City.

That some titanates, especially $BaTiO_3$ and barium-strontium titanate mixtures, exhibited unusual dielectric properties, was first disclosed in the U.S.A. by Wainer and Solomon in 1942 and 1943 in reports of the

Titanium Alloy Manufacturing Company. Our own work during 1943 and the spring of 1944, established the accurate dielectric characteristics of TiO_2 and of the alkaline earth titanates including that of barium-strontium mixtures, measured thermal expansion and temperature coefficients, correlated by X-ray studies the high peak in the dielectric constant of the barium-titanate and the barium-strontium materials with a polymorphic transition from pseudo-cubic to cubic and established that these ceramics represent a new class of ferro-electric materials. In addition to the " Curie " point at high temperature just mentioned, we found two regions of anomalous response at lower temperature in $BaTiO_3$, located at about + 10° C. and − 70° C. at which the dielectric constant at higher field strength shows pronounced peaks. Discontinuous regions of expansion from 80° to 120° C., and −20° C. to + 10° C., accompany the two higher regions of dielectric anomaly. It is very gratifying to learn from Miss Megaw's more precise X-ray studies that the pseudo-cubic phase is tetragonal and that apparently in the 10° C. region the axial ratio c/a of $BaTiO_3$ reaches its maximum.

After establishing the ferro-electric nature of $BaTiO_3$ and its mixtures with $SrTiO_3$, the temperature region from − 170° C. to + 150° C. was investigated with cathode ray oscillograms as shown in our second report, and then the non-linear dielectric properties of the materials were studied in more detail. It has been possible to generate piezo-electricity in the substances, to excite the ceramics to resonance vibrations and to obtain a much better insight in the dielectric mechanism involved. A part of this study, carried through by S. Roberts as a doctorate thesis in this laboratory, will be published shortly by him in the *Physical Rev.*

Mr. Whiffen (*Oxford*) (*communicated*): The high frequency loss of ice observed by Lamb (his Fig. 2) increases very rapidly with temperature. This rise may be due to a shift of the loss maximum towards 10 cm. wave-length or to an increase in the height of the loss curve.

If the maximum is at shorter wave-lengths than 10 cm., as is suggested by the absence of any indication of the second loss region in the high frequency tail of the loss curve in Fig. 4, then the first alternative requires that the peak of the absorption moves to longer wave-lengths with increase of temperature. The rate process would have a negative temperature coefficient, a most unlikely state of affairs.

The other conclusion, namely that the height of the loss curve increases, suggests that the number of molecules able to take part in the loss process increases with temperature. Very possibly there are small patches of ice that melt below 0° C. The loss frequency curve at one temperature would then resemble that of water and have a maximum at about 2 cm. wave-length. The magnitude of the loss is such that if 1/10,000th part of the ice were liquid at − 1° C. this would be sufficient to explain the observed value of $\tan \delta$.

A very small impurity in the ice would account for the existence of water below 0° C., since during the freezing process the impurities would be concentrated in the unfrozen liquid and would finally be present in sufficient local concentration to depress the freezing point of ice by 10-20° C. It would be interesting to see the effect of impurities added intentionally; a rough calculation shows that impurities to the very small extent of 1 mole/10^4 litres of the original water sample might account for Lamb's results.

Water is also able to remain liquid below 0° C. under pressure. A pressure might arise if small portions of water were entrapped by ice during the freezing process and were unable to expand freely as they tried to crystallise. It must be remembered that pressure may affect the frequency of the loss maximum so that the observed high frequency loss maximum in ice might differ a little in wave-length from that found for supercooled water at the same temperature either by direct measurement or by extrapolation from higher temperatures.

Contents

PAGE

General Introduction. By the President 1

A. Theoretical—

Dipolar Interaction. By H. Fröhlich 3

The Influence of Hindered Molecular Rotation on the Dielectric Polarisation of Polar Liquids. By John G. Kirkwood . 7

Theory of the Crystalline Field in Solid and Liquid Polar Dielectrics. By Edmond Bauer and Daniel Massignon 12

A Contribution to the Theory of the Internal Electrical Field. By C. J. F. Böttcher 16

A Comparison of Debye's and Onsager's Theories of the Dielectric Constants of Concentrated Polar Media. By F. C. Frank . 19

Some Mathematical Models Representing Polar Molecules in Crystals. By F. C. Frank 24

General Discussion.—Mr. D. M. Ritson, Dr. F. C. Frank, Prof. E. Bauer, Dr. C. J. F. Böttcher, Dr. H. Fröhlich . . . 36

B. Theory of Relaxation Times—

Experimental Curves of Losses and of Anomalous Dispersion as a Basis of a Spectral Method. By Pierre Girard and Paul Abadie 40

Power Factor and Temperature Coefficient of Solid (Amorphous) Dielectrics. By M. Gevers and F. K. du Pré 47

The Distribution of Relaxation Times in Dielectrics. By C. G. Garton 56

On the Frequency Dependence of the Dielectric Constant in Dipolar Solids. By R. A. Sack 61

A Note on the Analysis of Dielectric Measurements. By S. Whitehead 66

General Discussion.—Dr. F. C. Frank, Mr. C. G. Garton, Dr. M. Magat, Mr. D. H. Whiffen, Dr. S. Whitehead. 75

C. Breakdown—

Electric Breakdown of Solid Dielectrics. By A. von Hippel . 78

General Discussion.—Dr. H. Fröhlich, Dr. R. W. Lunt, Dr. S. Whitehead, Dr. A. E. W. Austen 87

D. Measurement Technique and Terminology—

The Representation of Dielectric Properties and the Principles Underlying their Measurement at Centimetre Wavelengths. By Willis Jackson 91

PAGE

Dielectric Absorption in Benzene and Liquid Paraffin Solutions at Ultra High Frequencies. By Willis Jackson and J. G. Powles 101

Some Measurements of the Permittivity and Power Factor of Low Loss Solids at 25,000 Mc./sec. High Frequency. By R. P. Penrose 108

Measurements on the Absorption of Micro-waves. Parts I-II. By D. H. Whiffen and H. W. Thompson 114

Dielectric Properties of Water. By C. H. Collie, D. M. Ritson and J. B. Hasted 129

Wave Guide Measurements of Dielectric Absorption of Solutions of Polar Substances in Non-polar Solvents. By H. W. Hall, I. G. Halliday, W. A. Johnson and S. Walker 136

Extension of the Measurements of Dispersion and Absorption by Liquids to the Region of Centimetric Radio-Electric Waves. By P. Abadie 143

Some Measurements on the Absorption of Centimetric Waves by Liquid Dielectrics. By F. J. Cripwell and G. B. B. M. Sutherland . 149

Note on the Comparison of Entropy and Energy Changes in Diffusion and Dielectric Relaxation Processes. By D. L. Levi 152

General Discussion.—Dr. R. P. Penrose, Mr. D. H. Whiffen, Mr. C. H. Collie, Dr. J. G. Powles and Prof. W. Jackson, Mr. H. W. Hall, and Mr. I. G. Halliday, Dr. J. Lamb, Dr. Magat, Dr. H. W. Thompson, Dr. G. Gee, Prof. E. Bauer, Dr. F. C. Frank, Dr. H. Pelzer, Mr. L. E. Sutton, Dr. R. J. W. Le Fevre, Dr. J. H. van Santen, Dr. J. Weiss, Dr. C. J. F. Böttcher, Dr. A. Schallamach, Mr. D. M. Ritson, Dr. Sutherland, and Mr. Cripwell 155

E. Applications, including Liquid Systems—

On the Application of Dielectric Measurements to Chemistry. By L. E. Sutton 170

Molecular Freedom in Solid Dielectrics. By Charles P. Smyth . 175

The Dielectric Relaxation of Mixtures of Dipolar Liquids. By Adolf Schallamach 180

The Radio-Frequency Absorption Spectra of Solutions of Electrolytes. By Jan Forman and Dennis J. Crisp 186

General Discussion.—Dr. M. Magat, Prof. E. Bauer, Mr. D. M. Ritson 193

F. Long Chain Compounds—

The Dielectric Properties of High Polymers. By R. B. Richards 194

The Dielectric Properties of Chlorinated Polythenes at Radio Frequencies. By W. G. Oakes and R. B. Richards . . 197

Dipole Orientation in Solutions of Esters in Polyisobutene and Polythene. By K. W. Plessner and R. B. Richards . . 206

PAGE

Dielectric Dispersion and Absorption in Natural Rubber, Neoprene, Butaprene NM and Butaprene S, Gum, and Tread Stocks. By W. C. Carter, M. Magat, W. C. Schneider and C. P. Smyth 213

General Discussion.—Dr. H. Fröhlich, Dr. A. Schallamach, Dr. R. F. Tuckett, Mr. L. Mullins, Mr. R. B. Richards, Dr. R. J. W. Le Févre 220

G. Solids—

Crystal Structure of Barium Titanium Oxide and other Double Oxides of the Perovskite Type. By Helen D. Megaw . . 224

The Permittivity of Polycrystals of the Perovskite Type. By D. F. Rushman and M. A. Strivens 231

Measurements of the Dielectric Properties of Ice. By J. Lamb . 238

General Discussion.—Dr. W. Reddish, Dr. W. Plessner, and Prof. W. Jackson, Prof. E. Bauer, Dr. W. H. Taylor, Mr. D. M. Ritson, Dr. G. H. Jonker, Dr. J. H. van Santen, Mr. M. A. Strivens, Dr. Rushman, Dr. F. C. Frank, Dr. H. Megaw, Prof. A. von Hippel, Dr. R. G. Brekenridge, and Mr. D. H. Whiffen 244

AUTHOR INDEX*

Abadie, P., 40, 143.
Austin, A. E. W., 89, 90.
Brauer, E., 12, 37, 162, 193, 246.
Böttcher, C. J. F., 16, 38, 165.
Brekenridge, R. G., 252.
Carter, W. C., 213.
Collie, C. H., 129, 156.
Cripwell, F. J., 149, 168.
Crisp, D. J., 186.
du Pré, F. K., 47.
Forman, J., 186.
Frank, F. C., 19, 24, 36, 37, 40, 75, 159, 163, 250.
Fröhlich, H., 3, 39, 87, 220, 223.
Garton, C. G., 56, 75.
Garner, W. E., 1.
Gee, G., 101.
Gevers, M., 47.
Girard, P., 40.
Hall, H.W., 136.
Halliday, I. G., 136.
Hasted, J. B., 129.
Jackson, W., 91, 99, 101, 156, 163, 244.
Johnson, W. A., 136.
Jonkers, G. H., 248.
Kirkwood, J. G., 7.
Lamb, J., 148, **238.**
Le Fevre, R. J. W., **162, 222.**
Levi, D. L., 152.
Lunt, R. W., 88.
Magat, M., 77, 160, 169, 170, 193, **213.**
Megaw [Miss] H. D., 225, 246, 252.
Massignon, D., 12.
Mullins, L., 221.
Oakes, W. G., 197.
Pelzer, H., 164.
Penrose, R. P., 108, 155, 159.
Plessner, K.W., 206, 244.
Powles, J. C., 101, 156, 157, 163.
Reddish, W., 244.
Richards, R. B., 194, 197, 206, 222, 223.
Ritson, D. M., 36, **129,** 168, 193, 246.
Rushman, D. F., 231, 249.
Sack, R. A., 61.
Schallamach, A., 166, **180,** 220.
Schneider, W. C., 213.
Smyth, C. P., 175, 213.
Strivens, M. A., 231, 249.
Sutherland, G. B. B. M., 149, 168.
Sutton, L. E., 165, **170.**
Taylor, W. H., 246.
Thompson, H. W., **114, 122,** 160, 166.
Tuckett, R. F., 221.
van Santen, T. H., 165, 248.
von Hippel, A., 78, 252.
Walker, S., 136.
Weiss, J., 165.
Whitehead, S., 66, 78, 89.
Whiffen, D. H., 77, **114, 122,** 156, 164, 166, 253.

* The references in heavy type indicate papers submitted for discussion.

The Faraday Society reserves the Copyright of all Communications published in the "Transactions".

Transactions of the Faraday Society.

FOUNDED 1903.

TO PROMOTE THE STUDY OF ELECTROCHEMISTRY, ELECTROMETALLURGY, CHEMICAL PHYSICS, METALLOGRAPHY, AND KINDRED SUBJECTS.

Vol. XLII B. 1946. 20s.

A GENERAL DISCUSSION

ON

SWELLING AND SHRINKING

(For Table of Contents see page 303)

GURNEY AND JACKSON
LONDON: 98 GREAT RUSSELL STREET
EDINBURGH: TWEEDDALE COURT

OFFICERS AND COUNCIL, 1946-1947.

President.

Prof. W. E. Garner, C.B.E., D.Sc., F.R.S.

Vice-Presidents who have held the Office of President.

Prof. F. G. Donnan, C.B.E., Ph.D., F.R.S.
Prof. C. H. Desch, D.Sc., F.R.S.
Prof. N. V. Sidgwick, Sc.D., D.Sc., F.R.S.
Prof. M. W. Travers, D.Sc., F.R.S.
Prof. E. K. Rideal, M.B.E., D.Sc., F.R.S.

Vice-Presidents.

Prof. A. Findlay, D.Sc.
Sir Charles Goodeve, D.Sc., F.R.S.
R. O. Griffith, D.Sc.
Sir John Lennard-Jones, K.B.E., D.Sc., F.R.S.
R. Lessing, Ph.D.
Prof. H. W. Melville, Ph.D., F.R.S.
Prof. M. Polanyi, Ph.D., M.D., F.R.S.

Honorary Treasurer.

R. E. Slade, D.Sc.

Chairman of the Publications Committee.

Prof. A. J. Allmand, D.Sc., F.R.S.

Ordinary Members of Council.

R. P. Bell, M.A., F.R.S.
G. M. Bennett, M.A., Sc.D., F.R.S.
A. Caress, Ph.D.
Prof. C. W. Davies, D.Sc.
G. Gee, Sc D.
G. S. Hartley, D.Sc.
F. I. G. Rawlins, M.Sc., F.R.S.E., F.Inst.P.
Sir Robert Robertson, K.B.E., D.Sc., F.R.S.
H. W. Thompson, M.A., D.Phil., F.R.S.
O. H. Wansbrough-Jones, O.B.E., Ph.D.

Honorary Secretary.

G. S. W. Marlow, B.Sc., F.R.I.C.

Editor.
F. C. Tompkins, Ph.D.

Assistant Secretary.
Miss B. E. Kornitzer.

6 Gray's Inn Square, London, W.C.1.
(Telephone: Chancery 8101.)

Publications Committee.

Prof. A. J. Allmand, D.Sc., F.R.S. (Chairman).

Prof. N. K. Adam, M.A., Sc.D., F.R.S.
Prof. W. T. Astbury, M.A., Sc.D., F.Inst.P., F.R.S.
T. Beacall, O.B.E.
C. A. Coulson, M.A., Ph.D., D.Sc.
Prof. H. J. Emeléus, D.Sc., F.R S.
Prof. M. G. Evans, D.Sc., F.R.S.
Prof. W. E. Garner, C.B.E., D.Sc., F.R.S.
G. Gee, Sc.D.
Prof. E. A. Guggenheim, M.A., Sc.D., F.R.S.
Prof. C. N. Hinshelwood, Sc.D., F.R.S.
Prof. W. C. M. Lewis, D.Sc., F.R.S.
Prof. H. W. Melville, Ph.D., F.R.S.
Prof. M. Polanyi, Ph.D., M.D., F.R.S.
Prof. J. T. Randall, D.Sc., F.Inst.P., F.R.S.
Prof. E. K. Rideal, M.B.E., D.Sc., F.R.S.
Sir Robert Robertson, K.B.E., D.Sc., F.R.S.
D. W. G. Style, Ph.D.
Prof. W. F. K. Wynne-Jones, D.Sc.

A GENERAL DISCUSSION ON SWELLING AND SHRINKING

A General Discussion on Swelling and Shrinking was held at the Royal Institution, London, from the 24th to the 26th September, 1946, the President, Professor W. E. Garner, F.R.S., occupying the Chair throughout. About 300 members and visitors were present, and among the distinguished overseas members, the President welcomed the following :—

Prof. T. Alfrey (Brooklyn), Dr. L. M. Arnett (U.S.A.), Dr. A. H. W. van Aten (Holland), Dr. C. E. Black (U.S.A.), Dr. and Mrs. C. J. F. Bottcher (Rotterdam), Dr. D. B. Clapp (U.S.A.), Lt. Walter Clarke (U.S.N.), Lt. Commander Daly (U.S.N.), Dr. H. Debruyn (Geleen), Dr. D. G. Dervichian (Paris), Dr. P. M. Doty (Notre Dame, Indiana), Prof. E. Fauré-Fremiet (Paris), Dr. G. E. van Gils (Holland), Monsieur and Madame S. Henin (Versailles), Dr. J. J. Hermans (Utrecht), Dr. P. H. Hermans (Utrecht), Dr. M. Magat (Paris), Dr. M. Mathieu (Paris), Dr. J. Mering (Paris), Dr. J. L. van der Minne (Delft), Dr. W. L. J. de Nie (Amsterdam), Dr. N. van Olphen (Amsterdam), Dr. J. Th. G. Oberbeek (Eindhoven), Prof. P. Putzeys (Louvain), Dr. W. E. Roseveare (U.S.A.), Dr. G. E. Rumscheidt (Amsterdam), Dr. L. S. Salisbury (U.S.A.), Dr. T. Salomon (Strasbourg), Dr. A. J. Staverman (Leiden), Prof. E. R. Theis (Lehigh University, U.S.A.), Dr. D. Vermaas (Utrecht), Dr. A. J. Wildschut (The Hague).

The papers to be discussed had all been issued in Advance Proofs, these, and the reports of the Discussion thereon appear in the following pages.

SWELLING AND SHRINKING

A GENERAL DISCUSSION HELD AT THE ROYAL INSTITUTION, 21 ALBEMARLE STREET, LONDON, W. 1 ON 24TH, 25TH AND 26TH SEPTEMBER, 1946.

GENERAL INTRODUCTION

By J. D. Bernal.

The discussion of this subject by the Faraday Society is particularly timely. It affords the first opportunity since the war for colloid scientists to meet together and for us to welcome our colleagues from abroad from whom we have been separated for such a long period. It is also well timed in relation to the development both of the subject itself and of its vastly increased range of applications which have arisen in the course of the war. In the study of colloidal phenomena, particularly of gels, we are at the beginning of a new period of advance, due largely to the introduction of new physical methods of examination, notably those of the electron microscope and X-ray analysis, and at the same time to the development of new theories on the interaction of particles and structures of sols and gels. These developments of methods coincide with, and are in part the consequence of, greatly increased technical utilisation of colloidal materials, particularly in the gel form.

In field after field of technology, colloidal properties are found to be of paramount importance. This is particularly so in plastics, both in sheet and fibre form, in natural and synthetic rubbers, and in the deeper study of natural fibres of technical or biological interest, such as wool and muscle. The development of soaps and detergents offers a new field for the study of gels, as does that of clays and other natural and industrial pastes.

The drive now is to relate the mechanical properties of these substances with their basic physico-chemical constitution. Most of the phenomena of industrial interest are outside the range of molecular chemistry, and can only be described in terms of micelles and long-range forces. It was my own interest in one of these systems, notably that of cement, that was in part responsible for the calling of the present meeting. At a meeting earlier this year of the Building Industries Group of the Society of Chemical Industry, the more technical aspects of this subject were discussed, but here we have the opportunity of attempting to deepen our understanding of the fundamental processes underlying the mechanics of gels, and particularly of their swelling and shrinking properties.

As one who has come only recently into the field of colloid science, I would like to say something about how this baffling and complex subject appears from the outside in terms of the more securely established fields of science. What is most immediately striking is the enormous amount of work that has been done in this field, but at the same time, with a few luminous exceptions such as the work of Freundlich, how extremely limited the researches have been to the properties of particular systems and how difficult therefore it is to obtain general perspectives. Here is a case which calls out for fundamental experimental research and theoretical probing which would well repay any diversion of effort from the study of the immediate applications.

In the first place the definition of gel is still a very loose one. It has been taken to cover any fluid-containing system showing mechanical rigidity, with a structure which cannot be elucidated by the optical microscope. It is extremely difficult to draw a line between a weak gel and an anomalous fluid showing viscous elasticity, or between a gel and a paste of microscopically visible particles. The other characteristic of gels which we are particularly concerned with at this conference is the relation of their structures to the fluid content, in other words their swelling and shrinking. Old models based on simple capillarity cannot explain these phenomena, nor can, as we shall hear later on in the conference, osmotic phenomena satisfactorily account for them. We are faced, therefore, with the need to look deeper into the structure of gels to see whether it contains the clues to their peculiar behaviour.

The use of the electron microscope provides one clue to the solution of this problem. In practically every case that has been examined so far, there has been evidence of definite structure, most often of crystalline structure, which is also revealed by X-rays. Sometimes this structure seems thoroughly confused and irregular, but more often it shows degrees of parallelism leading to the formation of tactoid structures. The only common feature of gels, as indeed of sols, has been that the particles concerned have at least one dimension of the order of between 10 and 1000 A. Another feature, however, though not so definitely proven, may be the one which characterises gels as distinct from sols, namely their containing rigid structures or agglomerates, in one or two dimensions, of much greater size, such as are provided by the fibre molecules or micellæ of natural or artificial fibres, by the ribbon-like micellæ of soaps or by the flat crystalline particles of hydroxide gels or clays.

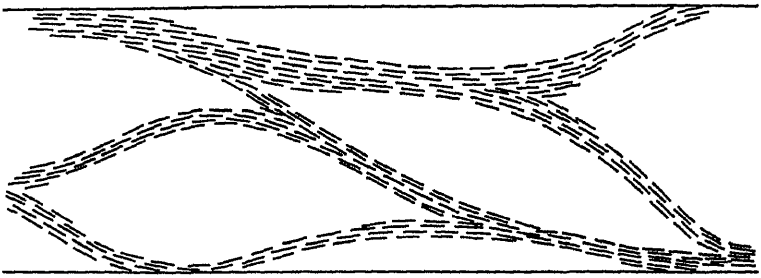
While these are common properties, the gels differ amongst each other so markedly that it would seem worth while to try and classify them according to the most reasonable available views of their internal structure. The most obvious division is that based on the relation of the solvent to the gel substance. If a gel consists of rigid particles and solvent only it may be called a simplex gel, such for instance as silica gel appears to be. On the other hand many gels have recently been shown to contain solvent in two forms, one closely bound to the elementary micellæ and separating them by distances which may range from 10 to 1000 A.; and the remainder, which fills much larger gaps between the meshes of the micellar network. Such for instance is the structure of the gels of tobacco mosaic virus, which have been studied both by X-rays and by the electron microscope, and that of soap gels and of bentonite gels. If there are two solvents, an aqueous and non-aqueous, these may be associated in different ways with the gel particles, forming gels that might properly be called complex, such as the mixtures of soap, water and hydrocarbons which are probably the bases of detergent action. In the case of duplex gels, it is not the association of solvent with the colloidal particles themselves that constitutes the gel structure. They may in fact form finite assemblages of a liquid crystalline character which usually appear as tactoids. It is only when tactoids fuse together, forming a tactoidal network (Fig. 1) that a true gel is formed.

In so far as this gel property depends on the existence of extended rigid particles in two dimensions, it will be formed whenever such extension is provided, either by the pre-existence of such particles as in natural fibres or clays, or by their formation in the body of the solvent. Meyer[1] first pointed out the close analogy between the setting of a gel and crystallisation from a supersaturated solution. In fact many inorganic gels, such as those of vanadium pentoxide and ferric hydroxide, seem on electron microscope evidence to be formed in this way. Rapid crystallisation inevitably yields crystals much extended in the direction of the greatest crystallisation velocity. The greater the difference between crystallisation velocities in different crystal directions, the greater tendency to form gels. Particles in gels of this sort may be considered analogous to polymer fibres, the greatest crystallisation direction being the main polymer chain direction, and the slower directions of crystal growth corresponding to the lateral bonding of the polymer chains.

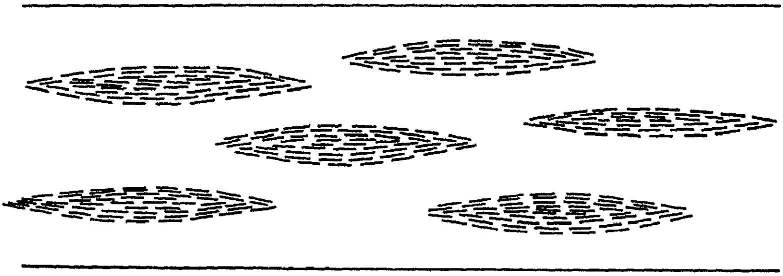
---

[1] Meyer and van der Wyk, *Helv. Chim. Acta*, 1937, **20**, 1331.

The great majority of gels are probably based on these crystal polymer particles, but there is at least one, and that an important one, which seems to be of a different nature, namely silica gel. Electron microscope photographs of silica gel indicate that it is formed from silica sol particles, themselves almost spherical in shape, stuck together to form tree-like branches (Fig. 2(*a*)). This type of packing of spheres differs very sharply from the close packing which occurs in the atoms of crystals or in the grains of sand (Fig. 2(*b*)), and is clearly due to the fact that when two spherical particles of the sol meet they stick together and cannot roll over one another. This hypothesis leads logically to a branched array with an average co-ordination number of 2, and one which can occupy about one-third of the volume filled by the tree-like assembly. The existence of this type of gel indicates that gels may be expected not only when chemical or crystallographic conditions tend to the formation of fibres or sheets (polymer gels), but also when any sol with adhesive particles coagulates (agglomerate gels).



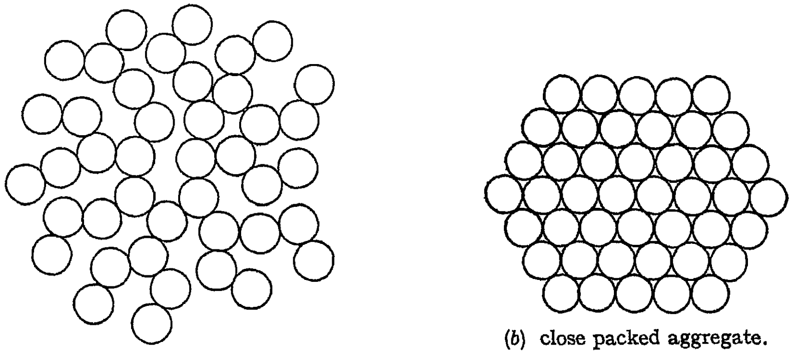
(*a*) in form of tactoids.



(*b*) in form of tactoidal network or duplex gel.

FIG. 1.—Alternate arrangements of long particles with inter-particle equilibria.

To explain the formation of the particles out of which a gel is built, however, clearly does not touch the distinction between gels and ordinary packings of macroscopic bodies. More and more evidence has recently been accumulating that between these extended particles and working through the liquid which bathes them, there are forces which act at distances of the same order as the size of the particles themselves. This was first shown by Zocher[2] in the case of the hydroxide gels showing Schiller structure, that is, periodicities of the order of the wavelength of visible light, 5000 A.; and later by Fankuchen[3] and myself in the gels of tobacco mosaic virus, where X-rays demonstrated conclusively regular hexagonal two-dimensional associations of elongated particles with distances apart ranging from contact up to at least 500 A. The regularity of the structures showed that we are here dealing with energy systems possessing minima at these distances, or to put it in another way, with attractive and repulsive forces in equilibrium, this equilibrium itself being determined by the ionic constitutions of the medium. Other work has shown that also in soaps, clays and inside the crystals of proteins, long-range forces, leading to equilibria, are involved. One of the most interesting features of this conference should be the discussion of the physical nature of these long-range forces.



(a) sticky particles forming tree of average co-ordination number two.

(b) close packed aggregate.

FIG. 2.—Two-dimensional scheme of aggregates of particles.

I am myself somewhat uncertain whether we are dealing here with a single or a complex phenomenon. It may well be that the forces working over a distance of the order of 10-30 A., such as occur in clays, soaps and proteins, may be of a different kind from the longer range forces observed in the virus and hydroxide tactoids; and that the first may in fact be fairly normal association through water or solvent molecules held in fixed positions. But whether this is so or not, the longer range forces need a different type of explanation. All attempts at explanation provided so far are based on interaction of ionic atmospheres, and possibly London forces between the particles themselves. But it is clear from the divergencies between the various calculations that the effect, being small, is one of higher order, and may require more refined theories than are at present available.

There is also another point which is difficult to explain on any theory proposed hitherto. Most gels studied in detail have been water gels, but gels are known, including those based on polyvalent ion soaps, which will be discussed in this conference, which are in hydrocarbon media where ions play no part, and yet these also seem to show phenomena of the same kind as those given by the ionic gels. It is true that hitherto no binary systems which would lead in theory to the existence of either phenomenon without the other, such systems will never be realised in practice. The large contribution of real velocity to the overall diffusion in the swelling process is directly dependent on a great discrepancy of molecular size, and is thus very great only in a polymer-solvent system. The great variation of diffusion coefficient with concentration, of the form described, is, however, due to a combination of causes, some of which can operate in other systems. One of these, and perhaps quantitatively the most important, is great variation from ideal solution thermodynamics in the direction of what Hildebrand describes as " negative deviation from Raoult's law ". This alone will cause the diffusion coefficient to go through a maximum,[9] as in the case of acetone and chloroform.[10] Only, however, in high polymer-solvent systems, where this type of deviation is extreme,[11] is this effect likely to be large enough. The other factors are the direct effect of discrepancy of size already referred to and the great viscosity variation which is responsible for the fall of diffusion coefficient in high concentration of polymer.

---

[2] Zocher, *Z. anorg. Chem.*, 1925, **147**, 91; Zocher and Heller, *ibid.*, 1930, **186**, 75; Bergmann, Löw-Beer and Zocher, *Z. Physikal. Chem.*, *A*, 1938, **181**, 301.

[3] Bernal and Fankuchen, *J. Gen. Physiol.*, 1941, **25**, 111.

A scale of the type we have described, while measuring real flow, will be uninfluenced by diffusion directly, the particles of indicator substance being too large to show appreciable diffusion. A given square of the scale will not necessarily always contain within its projection the same amount of polymer substance, since there may have been an appreciable change of concentration by pure diffusion as distinct from real flow. It we could use such a scale to give the x units, combined with values of concentration independently determined, in the basic diffusion equation, we should now measure two diffusion coefficients not simply related to each other. It is for this reason that we made the qualification, when saying that the two diffusion coefficients are not independent, that this was only true when the section of reference was fixed in a way not dependent on the diffusion process. In effect we upset this generalisation by using the scale because in so doing we have introduced a third (non-diffusible) component. In such a system, the contours of equally probable displacement in a given time interval δt of the c.g.s of different molecules are not, as in the case of interdiffusion of isotopes, concentric spheres. The unsymmetrical contours could be resolved into spherical components of radii proportional to $(\delta t)^{\frac{1}{2}}$ and a linear component proportional to δt. The linear component would be the same for all kinds of molecules and macroparticles and measures the real velocity.

We may note that, in the case where the intrinsic diffusion velocity (that measured, e.g., by reference to a non-diffusing component) varies with concentration from any cause, this variation will itself contribute a real velocity proportional to $\frac{\partial D}{\partial x} = \frac{dD}{dc} \cdot \frac{\partial c}{\partial x}$, due to the asymmetry of mean displacement by Brownian movement which again can be resolved into a real velocity component and a symmetrical random displacement.

The author is indebted to Dr. Conmar Robinson and Mr. J. Crank for much useful discussion of this subject, and to the directors of Messrs. Courtaulds, Ltd., for permission to publish this paper.

Courtaulds Ltd., Research Laboratory,
Maidenhead.

---

[9] Powell and Eyring, *Advances in Colloid Science*, Vol. I (Interscience, 1942) p. 208.

[10] Lemonde, *Ann. Physique*, 1938, 9, 539.

[11] Cf Gee and Treloar, *Trans. Faraday Soc.*, 1942, 38, 147.

DIFFUSION AND SWELLING OF HIGH POLYMERS.

PART II. THE ORIENTATION OF POLYMER MOLECULES WHICH ACCOMPANIES UNIDIRECTIONAL DIFFUSION.

By Conmar Robinson.

Received 2nd August, 1946.

It was explained in Part I that the swelling of a polymer will, at least in part, be due to a real velocity of the whole system as distinct from a true diffusion. One result of this is that when a solvent diffuses into a piece of polymer, restricted so as to allow swelling in one direction only, the polymer will move outwards in the region between the inner and outer boundaries of the swollen region with a velocity which increases with increasing distance from the inner boundary.

The polymer will in fact be stretched, and it may therefore be expected that the polymer molecules will be partly oriented and extended in the direction of diffusion. The process will be equivalent to the orientation by mechanical stretching in the swollen state which has been quantitatively investigated by Kratky and colleagues.[1]

In order to estimate the orientation which would be expected on the basis of previous work on mechanical stretching, we must estimate the extent of stretching at all points in the swollen polymer. Neglecting any small changes of volume which occur on mixing, it might be considered that the ratio of initial to present volume concentration of the polymer would be equal to the ratio of present to initial length. This, however, would only be true if the dilution of polymer was produced solely by variation of its real velocity of expansion. In fact, however, the pure diffusion of the polymer will have contributed to the dilution. To obtain a correct result we may make use of the fact that, as explained in Part I, inert particles suspended in the polymer which are themselves too large to show any appreciable diffusion, will have the same real velocity as the system as a whole. If such particles are incorporated in the polymer in the form of a scale, the (non-uniform) expansion of this scale on swelling will measure the degree of stretching of the polymer. Measurement of the extension will determine the maximum degree of orientation which might be expected to be reached at any point in the gel. Actually, as the solvent advances into the swelling polymer and progressively dilutes it, relaxation of the orientation by rotational diffusion will become increasingly important, so that the observed double refraction will fall increasingly below that expected from the degree of swelling.

As the conception of a real velocity of the whole system may be considered controversial, it may be pointed out that it is probably true that there will be only a small difference in the values for the expansion calculated from concentration variation (not yet determined), on the one hand, and from the swelling scale on the other, in the region where an appreciable degree of orientation can be observed in the experiments to be described.

In the following brief account of some preliminary experiments some of the qualitative aspects of the double refraction that accompanies the diffusion as the result of this orientation will be described. The experiments were carried out in suitably constructed cells placed on the stage of a polarisation microscope.

---

[1] Breuer, Kratky and Saito, *Kolloid Z.*, 1937, 80, 139.

Preparation of Diffusion Cells.—In all the cells used the piece of cellulose acetate sheet or foil of the required shape and size was mounted on one microscope slide and covered with a second. Swelling was restricted to one direction only by glass walls made from strips of glass of the same thickness as the acetate. In the earlier experiments the cells were cemented together, but later it was found more satisfactory to avoid cement by using a brass screw clamp at either end of the cell. The required thickness of cellulose acetate was obtained either by machining down a piece of sheet or by clamping together two or three pieces of foil. In experiments where the solvent was allowed to diffuse right through the material, the acetate formed the cross-piece of an H-shaped cell. In other cases, where diffusion did not reach the far edge, the walls of the cells, forming three sides of a rectangle, could conveniently be made of the cellulose acetate; this arrangement allowed diffusion to be compared while taking place simultaneously both at right angles and parallel to a given initial orientation of the sheet.

Material.—Two forms of commercial cellulose acetate were used in the experiment. The first was sheet, 3 mm. thick, containing 30 % plasticiser (tri-aryl-phosphate and di-alkyl phthalate mixtures), made by what is essentially a moulding process. The second was foil (0·25 mm. thick) containing 20 % plasticer, made by spreading of solution and drying. For some experiments the plasticiser was removed (as determined by loss in weight on drying) by soaking in a mixture of methyl ethyl ketone and butyl acetate (3 : 2 by volume). This mixture was also used for swelling the film prior to stretching where a chosen degree of *initial orientation* was produced.

The diffusing liquid was either acetone or methyl ethyl ketone. Since the cellulose acetate swells without dissolving in the second liquid, it is more suitable for studying the movement of the outer boundary than is acetone, but during the initial stages of the diffusion it gave rise to a cloudiness which made observation of the double refraction difficult. The photographs which are shown here were in all cases of experiments with acetone.

Optical Arrangement.—The cell was observed with a polarisation microscope using crossed nicols and with a first order red plate inserted at 45°. The microscope, as well as having the usual slit above the objective through which the first order red plate was inserted, had a slit in the eyepiece through which a quartz wedge could be inserted at right angles to the first order red plate. The stage of the microscope was oriented so that the edge of the cellulose acetate sheet which was to be observed was parallel to the length of the quartz wedge and at 45° to the crossed nicols. The advantage of using the first order red plate rather than only the crossed nicols were firstly that, with isotropic material in the field at the beginning of an experiment, the whole field, instead of being dark, was uniformly illuminated with red light, and secondly, that this tint, sometimes referred to as the sensitive tint, is particularly sensitive to small changes in the double refraction. For most of the observations a 1-inch objective and 6 X-eyepiece were used. A Hutchinson scale, with faint and dark lines, as in graph paper, was fitted into the eyepiece and oriented so that one set of lines ran parallel to the direction of insertion of the quartz wedge.

Description of Experiments.

Between crossed nicols and with the first order red plate in position before filling the cell with solvent the field appeared uniformly red or, if the cellulose acetate used had some degree of initial orientation, this showed a different, but uniform colour. Shortly after filling the cell, a sharp boundary appeared advancing into the cellulose acetate and parallel to the surface entered by the liquid. Behind this boundary, a " rainbow " of colour appeared with the bands of colour also parallel to this surface, the colours broadening out as the boundary advanced. The double refraction corresponding to the particular colour seen at a given distance from the edge could be estimated approximately by comparison with a polarisation colour scale such as that supplied by Messrs. James Swift. In this way it could be seen that highest double refraction occurred just behind the advancing front (e.g. with a thickness of 1·5 mm. of cellulose acetate a shift from first order red to third order green could be observed, corresponding to a retardation of about 1250 mμ).

The variation in the degree of double refraction in the direction of diffusion can be observed with great ease if the quartz wedge is inserted with its length parallel to the edge of the acetate. When the wedge is inserted with no double refracting material in the field a straight black line is seen, corresponding to that thickness of the wedge which causes extinction with the first order red plate. When a sheet of cellulose acetate of uniform double refraction is introduced into the field the black line will be displaced where it crosses the acetate parallel to its original position by a distance up or down the wedge corresponding to the sign and magnitude of the double refraction of the sheet. When the diffusion takes place there is a gradient of double refraction in the acetate at right angles to that in the wedge, the combined effect of which is to displace the black line in such a way as to give what appears as a graph seen against the Hutchinson scale, the ordinates of which measure the double refraction and the abscissae the distance diffused.

On moving the acetate parallel to its edge the curves remain stationary and unchanged, as would be expected from the straightness of the rainbow. The image of both scale and curve were readily photographed and in some cases the whole course of an experiment was recorded by cinematography, using either white light with a Kodachrome film in order to record the interference colours or ordinary film with a high pressure mercury vapour lamp and a Wratten No. 62 green filter where monochromatic light was required. In the latter case the quartz wedge, instead of showing one black line and a series of colour bands, shows a series of black lines corresponding to first, second and third order, etc., and since the distance between two of the lines corresponds to a shift of one wave-length (546 $m\mu$) this distance can be used as a scale of reference for estimating the double refraction at any point on the curve. Thus in Fig. 1*a* the peak of the curve corresponds to a retardation of 0·85 of a wave-length or 465 $m\mu$.

Fig 1*a* and 1*d* illustrate successive positions of boundary and curve as they proceed across the cellulose acetate. AB is the inward moving boundary (the acetone is entering from the right) and CD is the dry edge of the cellulose acetate. The original position of the black line produced by the quartz wedge was collinear with EF, its position against the empty field. Immediately to the right of the inward advancing boundary the line is displaced upwards during diffusion (see dotted line) but the whole of its position is not shown in the drawing as it was for some distance almost vertical and coincident with the position of the boundary. It will be seen that the height of the peak of the curve (maximum double refraction) remains constant or very nearly constant during the diffusion and proceeds just behind the advancing inner boundary.

On the other hand, at any point in the wake of this peak, the double refraction decreases as the boundary advances until the acetate at that point has fully relaxed and become isotropic (i.e., the black line tends to become collinear with its position against the empty field).

This tendency of the curve to flatten out behind the advancing peak, as shown in Fig. 1, is due to the rotational diffusion destroying the orientation as the cellulose acetate becomes progressively diluted by solvent or swelling liquid.

The double refraction produced was found to be proportional to the thickness of the cellulose acetate used. This would not have been the case if the double refraction had been caused by a shearing effect due to distribution of velocity of flow instead of being a direct consequence of the stretching process.

To interpret the double refraction curve fully in terms of orientation, it would be necessary to determine the variation of concentration with distance in the direction of diffusion. This has not yet been done. In Fig. 2 (Plate I), however, a photograph is shown of diffusion into two pieces of acetate film placed one on top of the other. On the upper surface of the lower piece a scale has been ruled with lines 0·005 inch apart. One set of

PLATE I

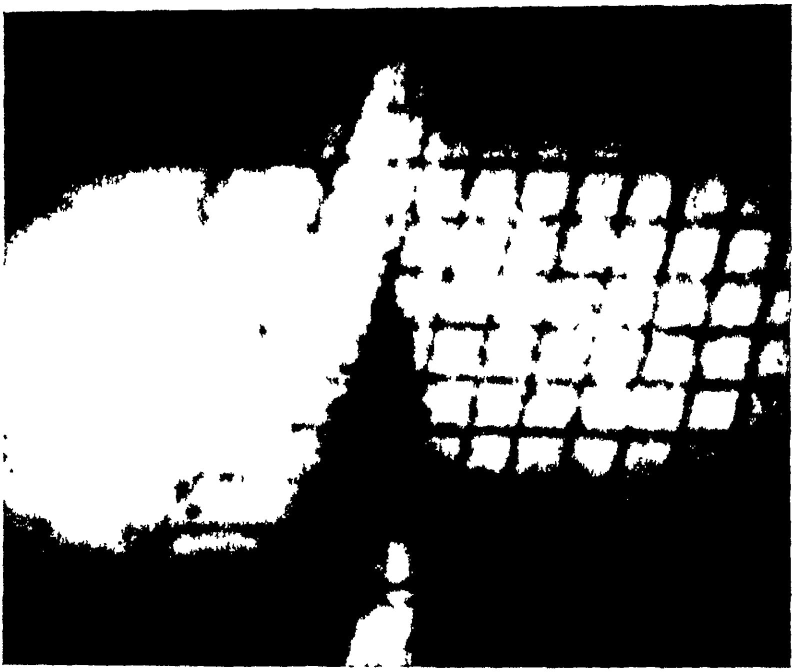


FIG 2.—Expanding scale and quartz wedge curve. (Acetone entering from left.)

[To face page 14

lines are inclined at an angle of $\tan^{-1}0{\cdot}2$ to the normal to the other set which lie in the direction of diffusion. This enables a vernier principle to be applied to the relation between swelling and distance. The expansion of the scale can be clearly seen in the wake of the advancing peak, some of the "squares" having extended to at least three times their original length. It is at once seen that the maximum orientation does not occur where the swelling is greatest, since here dilution has allowed relaxation to take place. It should be noted that the observed double refraction would not be proportional simply to orientation, since the doubly refracting material is being diluted at the same time as its intrinsic double refraction is increasing. On the basis of Kratky's results we should expect a maximum retardation in film initially 0·25 mm. thick of about 0·6 of a wave-length

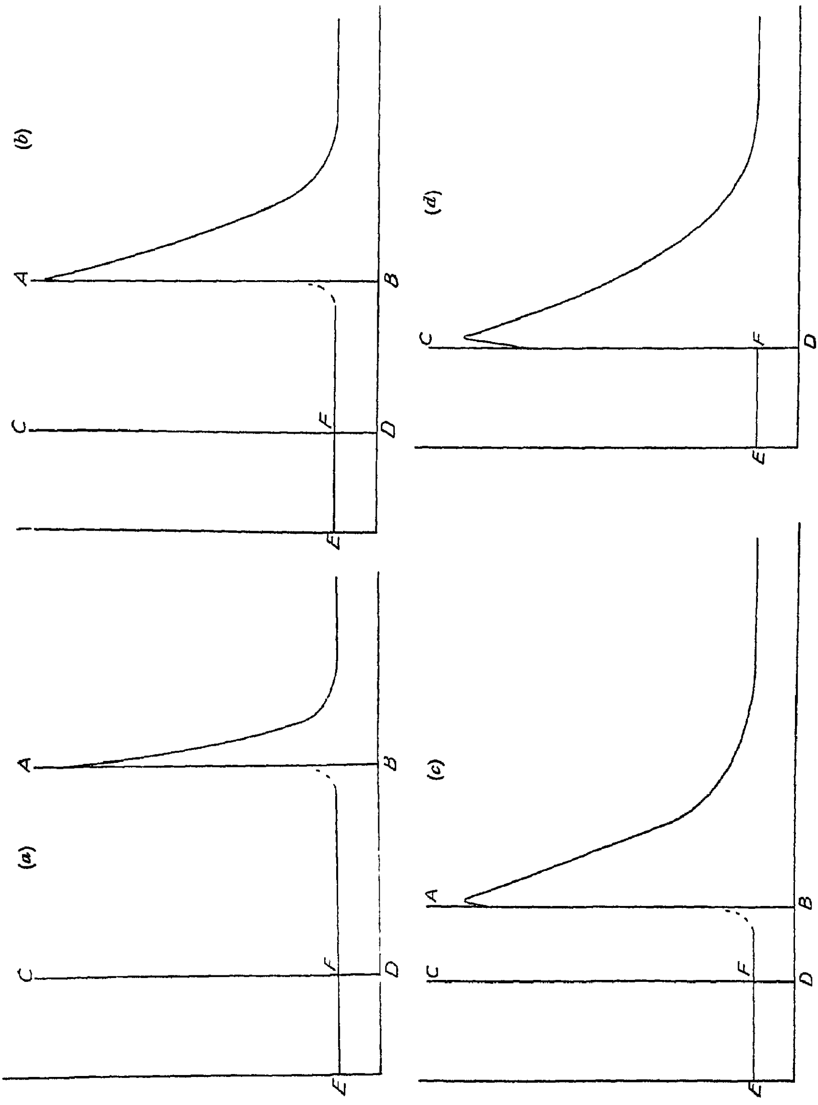


FIG. 1(*a*) to 1(*d*).—Double refraction accompanying unidirectional diffusion of acetone into cellulose acetone. (Acetone entering from the right.)

for green light which would be attained at an extension to about twice the original length. The peak we observe, however, is much sharper than would be expected from this cause, and appears at about 12 % extension beyond original length. Its height corresponds to a retardation of 0·3 of a wave-length for 0·25 mm. film which is actually higher than would be expected (0·2 wave-length) from Kratky's results. Our own observations, however, indicate that it is not higher than the double refraction which results from stretching our acetate mechanically by 12 % when swollen in the methyl ethyl ketone-butyl acetate mixture. Moreover, the effect of acetone on double refraction at a given degree of orientation has not been separately examined.

Fig. 3*a* to 3*c* show photographs of diffusion into pieces of stretched sheet, the diffusion taking place parallel, at right angles and at 45° respectively to the direction of stretching. The photographs illustrate the effect of different types of initial orientation on the double refraction curve. If the initial orientation had no effect on the curve, the peak would be expected to be the same as that in 3*c* in all three cases. Actually the peaks of 3*a* and 3*b* are intermediate between this value and those which would be expected if the effect of diffusion and initial orientation were purely additive.

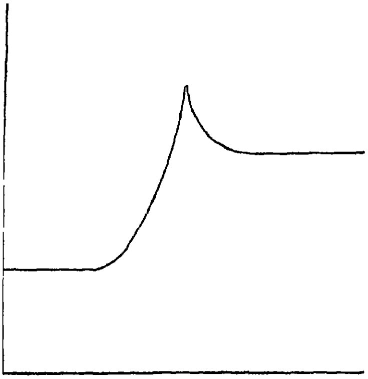


FIG. 3(*a*).—Direction of diffusion parallel to direction of stretching.

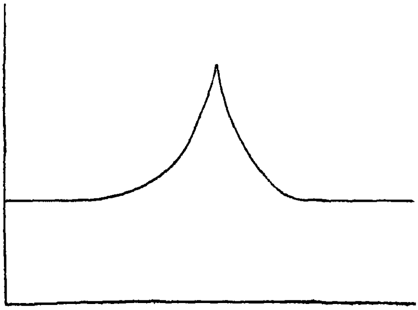


FIG 3(*b*) —Direction of diffusion at 45° to direction of stretching.

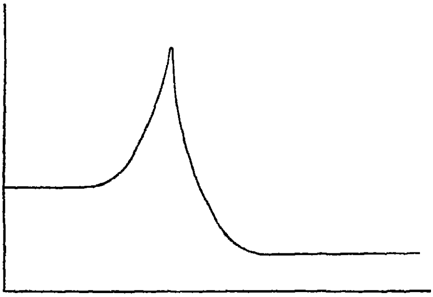


FIG. 3(*c*).—Direction of diffusion at right angles to direction of stretching.

FIG. 3(*a*) to 3(*c*).—Dependence of the shape of the double refraction curve on initial orientation in the cellulose acetate. (Acetone entering from the left.)

It will be noticed that the curves shown in Fig. 3 are not as vertical as those in Fig. 1 on the dry side of the boundary. It was at first thought that that this might be due to the removal of plasticiser in the film used in the experiments of Fig. 3. However, an experiment using unstretched plasticised film gave the same shaped peak, quite unlike those obtained with the *sheet* used in the experiment of Fig. 1. The reason for the difference has not yet been found, but it should be noted that in the film there is a considerable degree of non-selective uniplanar orientation (Sisson),[2] not detectable in light transversing normally, whereas the sheet is, except in one structure, more nearly isotropic.

If diffusion is allowed to take place from both sides of a piece of acetate, two advancing peaks can be seen approaching one another. On meeting and coalescing, the height of the peaks decrease until, after a fairly short time, the material becomes completely isotropic as would be expected from the symmetry of the system.

If, on the other hand, a piece of acetate into which acetone had been allowed to diffuse from one side (without penetrating to the other) is allowed to dry, a peak is produced *below* the original position of the black line, with the optical arrangement used, corresponding to a *negative* orientation in the direction of diffusion (that is to say a tendency of the molecules to avoid this direction, which is the " non-selective uniplanar orientation " of Sisson). Under suitable conditions, both orientations can be observed simultaneously in the same piece of acetate which then shows an S-shaped curve with the quartz wedge.

The magnitudes of the reversed peaks obtained on drying showed a greater degree of double refraction than those obtained when the liquid was entering, and in cells which were left empty this high degree of double refraction persisted for weeks and would presumably be permanent on account of the elimination of rotational diffusion. Few observations during drying have been made. They are much more difficult experimentally on account of the erratic adhesion of the shrinking cellulose acetate to both glass walls. Shrinking, in contrast to expansion, is not easily limited.

The author's thanks are due to Dr. G. S. Hartley for many discussions and suggestions, to Mr. C. D. Smith for technical assistance, and to Mr. R. McV. Weston for the cinematography. We also wish to thank the directors of Messrs. Courtaulds, Ltd., for permission to publish this research.

Messrs. Courtaulds' Research Laboratory,
Maidenhead.

---

[2] Chapter in Ott, *Cellulose and Cellulose Derivatives* (Interscience, 1943), p. 231.

THE REACTIONS OCCURRING IN THE GELATINISATION OF THE NITROCELLULOSES.

By Thérèse Petitpas and Marcel Mathieu.

Received in French, 31st August, 1946, and translated by Dr. F. D. Miles.

Investigations have been made to decide the action of different solvents on the nitrocelluloses. To a certain extent the nitrocelluloses can be considered as examples of polymer types with which a great variety of reactions can be readily carried out because of the stability of the cellulose chain and of the peculiar reactivity either of the —OH group of the cellulose molecule, or of the —ONO_2 group of the cellulose nitrates. In studying the behaviours of cellulose and of nitrocellulose, we have been able to formulate the probable mechanism of numerous reactions between the polymeric substances and their solvents or plasticisers.

It should be remembered that we have always tried to consider a given reaction in several different ways. Estimates of the absorption which is a function of such thermodynamic conditions as temperature and activity of the reacting substance have made it possible to define the *composition of the phase* containing the cellulose compound under all conditions.[1] X-ray diagrams [2] have allowed us to follow *the variations of structure* of the cellulose derivatives when solvent is absorbed. The essential precaution necessary here is to make X-ray diagrams only from such phases as are under well-defined thermodynamic conditions and are of definite composition. From calorimetric measurements all the *heat effects* accompanying the absorptions could be calculated.[3] Estimations of *birefringence* made with the polarising microscope and examinations of the *mechanical properties* [5] have led to the characterisation of the plastic phases which result from the absorption of solvent or plasticiser into the fibres of nitrocellulose. Finally, from *density measurements* we were able to say definitely what was the structure of a plastic phase obtained during the course of the absorption.

The details of this work have been published and it is unnecessary to repeat them here. We shall deal in this paper only with the essential conclusions which can be drawn from the investigations considered as a whole.

The first important observation which may be made is that solvents for nitrocellulose can be divided into two classes (1) those which in being absorbed modify the crystalline structure which X-rays reveal, and (2) those which cause no such modification. To the first, or as it may be called, the " acetone " type, belong the ketones—acetone, cyclohexanone, cyclopentanone and the alkyl acetates—methyl, ethyl, propyl, butyl, amyl. Camphor also belongs to this first class. To the second—which we may call the " methyl nitrate " type—belong the alkyl nitrates—methyl, ethyl, propyl, butyl, and very probably the nitrate of ethylene glycol and nitroglycerine. The absorption of water by cellulose, which takes place without changing the X-ray structure, must also be of this second kind.

We may recall the widely accepted idea of the structure of a fibre of cellulose, and more generally of a fibrous polymer. The degree of polymerisation of the cellulose chains may be from 3000 to 4000. These chains are arranged regularly in small crystalline domains but only in respect of certain parts of their length. Such domains of regular arrangements may contain about a thousand chains. They are grouped side by side, although with a certain variation from parallelism, and constitute in this way the crystallite, or elementary crystal. The crystallites are thus characterised by their mosaic structure. They are assembled to form the fibril with cylindrical symmetry, but without any other particular orientation, and the elementary fibre thus consists of fibrils in juxtaposition. In this way one can account for the degree of disorder which appears when one passes from the elementary chain to the macroscopic fibre. If instead of taking the fibre in cross-sections one considers it in length, there is seen to be another kind of disorder arising from the fact that the chain is formed by 3000-4000 residues united together, and that the regularity of arrangement does not extend over more than 300 or 400 residues in combination. A chain thus belongs to *several ordered elements* and serves as a bond between twenty or more of such regular elements. Between such elements the chain lies in a zone in which it is contiguous with other chains, but without any crystalline order. Schemes of such arrangements as this have frequently been given, in particular by Kratky for cellulose,[6] and by Bunn and Alcock [7] for polythene.

---

[1] Desmaroux, *Mémorial Poudres*, 1923, **20**, 1928, **23**. Mathieu, *Gelatinisation des Nitrocelluloses* (Hermans et Cie, 1936; Thèse, Paris, 1936). Petitpas, *Memorial Services Chim. de l'Etat* (Thèse, Paris, 1943).

[2] Mathieu, *loc. cit.* Petitpas, *loc. cit.*

[3] Calvert *et al.*, *Compt. Rend.*, 1941, **212**, 542; **213**, 126; 1942, **215**, 138; 1943, **217**, 482.

[4] Perrier and Mathieu, *ibid.*, 1942, **214**, 70.

[5] Raison and Mathieu, *ibid.*, 1941, **212**, 157.

[6] Kratky, *Kolloid Z.*, 1941, **96**, 301.

[7] Bunn and Alcock, *Trans. Faraday Soc.*, 1945.

We have therefore what may be called a " crystal " of cellulose material affected by disorder of two kinds (1) mosaic structure as in all crystals—this is probably very pronounced—and (2) domains in which the molecular chains are disposed without regularity.

The diagrams obtained with X-rays from cellulose or nitrocellulose are produced by the lattice arrangement existing in the zones of regularity. In so far as we are concerned with what have been called " topochemical " reactions, i.e. reactions in which the fibrous state is preserved, we can affirm that if the X-ray diagram is modified in the course of a reaction, it is because *the reactant has penetrated into the regular zones* and affected the grouping within them. When, on the other hand, the diagram is not changed, it is because the effect of *the reactant has been confined to the disordered zones and to the surface of the crystalline domains.*

The direct consequence of disorder in the structure of the fibres is therefore, from the chemical point of view, that the active groups (OH groups in cellulose ; O—NO_2 group in nitrocellulose) do not all play the same role in the material of the cellulose derivative. Some groups are enclosed and more or less blocked-up in the interior of the small domains of regularity ; some are at the surface of these domains ; others are attached to the half-isolated chains that constitute the zones of disorder. The OH groups on the surface, and those in the disordered regions are by no means completely free, because the small crystalline domains and the isolated portions adhere to each other by virtue of the same force that, in the zones of regularity, holds the elementary chains together in a regular lattice arrangement ; this is equally true for the parts of the chains which make up the disordered zones.

If the final consequence of the action of the reagent is that the cellulose material passes into solution, it follows that, for the first type of change already referred to, a " molecular " solution results by separation of the chain molecules ; for the second type, the dispersion does not go farther than the micellar stage. The type of solution resulting from the final action of the solvent cannot, however, be taken as the criterion of a reaction, because energetic action by a solvent on the surface of the crystalline regions can cause the chain to be torn off from the surface, and thus lead to molecular solution. It may be added that most solutions are formed from crystals and a solvent in this way, and that it is in this last case that the phenomenon of saturation, corresponding to a discontinuity, is observed.

The distinction between zones of order and zones of disorder in the structure of natural fibres has long been made, generally by referring to " crystalline " and " amorphous " phases. It seems, however, more precise and correct to speak of " zones " of order and disorder, because the word " phase " has a very definite thermodynamic signification. Another reason is that in all the reactions between fibres of cellulose compounds and solvents the fibres behave as a single phase.

In the same way, with regard to the action of numerous substances on the structure of cellulose derivatives, distinction has been made between reactions which are intracrystalline and these which are limited to the zones of disorder (also between " gaseous OH " and " crystalline OH ") and very often different authors have given good reason for localising the actions of a reagent in various zones. Thus Lieser and his collaborators [8, 10] consider that soda does not penetrate into the zones of regularity while Kratky [7] maintains that even these zones are swollen by the actions of the molecules of soda and water.*

We believe that systematic studies of absorption, the examination of

---

[8] Leiser and Fichtner, *Annalen*, 1941, 548, 195.

[9] Lieser and Jaks, *ibid.*, 1941, 548, 204.

[10] Leiser, Jaks and Glitscher, *ibid*, 1941, 548, 212.

* For recognition of the types of reactions of cellulose, see the important contribution of P. H. Hermans.

X-ray diagrams and determinations of density should permit us to characterise the type of reaction which is under consideration and to decide with fair certainty whether the reagent is acting throughout the mass of the fibres or whether it is confined to the regions of disorder, i.e. to the surfaces of the mosaic domains and to the irregular assemblies of molecules. We shall see that it is further possible to estimate the proportion of the total number of active groups which is to be ascribed to the regions of disorder.

I. Curves of Absorption.

We have seen that, from X-ray diagrams, we can tell whether a reagent penetrates into the structures of the crystalline zones of the fibre. Let us now examine the difference between the absorption curves obtained when the reagent remains in the zones of irregularity and those resulting from penetrating into the mosaic domains. The curves represent the molecular concentration N_2 of the solvent entering the solid phase as ordinates, and the activity a of the solvent as abscissae. The concentration N_2 is based on the group C_6 and is defined by $N_2 = \frac{N_s}{N_s + N_6}$, in which N_s is the number of molecules of solvent absorbed by a number N_6 of C_6 groups.

Ketones.—The absorption curves of ketones are reproduced in Fig. 1.

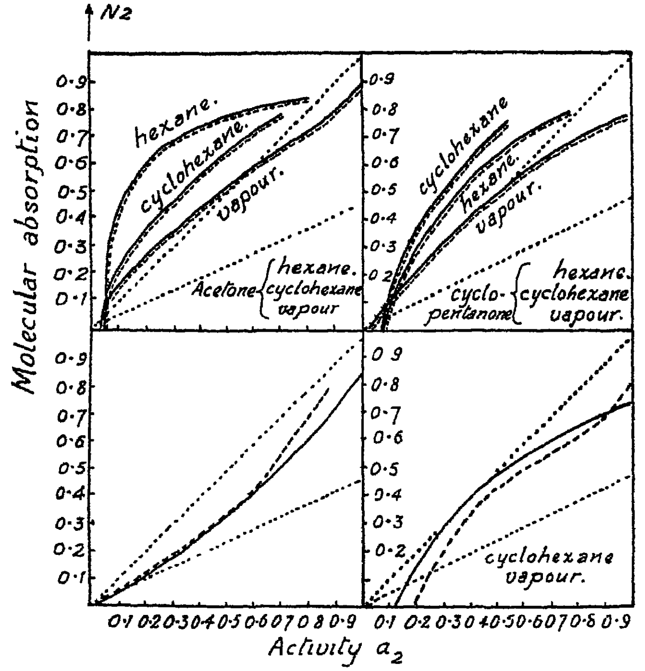


Fig. 1.

When extrapolated, they tend towards the point $N_2 = 1$ for $a = 1$, corresponding to infinite dilution.

Furthermore, having defined the absorption N_2 as above, curves are obtained for the vapours of ketones which are not far removed from the straight line at 45° to the axes, *such as would be obtained for an ideal solution.* In the case of the ketones, which represents the first type of reaction, the absorption can be taken to occur *in the whole mass* of the fibres. The regions of regularity are affected in the same way as those of disorder, or as the surface of the micelles; the facility of absorption is the same and the cohesion of the lattice, where it exists, offers only a negligible resistance to the fixation of the solvent.

Alcohol Nitrates.—With methyl nitrate, representing the second type of gelatiniser, the absorption curve of the vapour *is almost a straight line which, on extrapolation, tends to the point* $a = 1$, $N_2 = \frac{1}{2}$ (Fig. 2) and not to the point $a = 1$, $N_2 = 1$.

From the fact that the structure revealed by X-ray remains unchanged, we can legitimately conclude that only the regions of irregularity are affected by the penetration of solvent. On the other hand we can suppose, with much plausibility, that the absorption into the disordered regions is the consequence of forces and affinities of the same nature as those which cause the absorption into the whole mass of the fibres; in the case above, the absorption is limited to the regions of disorder only because the regular arrangement of the lattice cannot be disorganised by the solvent. Assuming this, the comparison of the absorption curves of the two solvent types leads directly to an evaluation of the number of C_6 groups not united in a regular lattice. *This number will be about half the total numbers of C_6 groups which constitute the cellulose material.* This factor of one-half is what has been assumed by all those authors who have compared the reactivity of the —OH groups situated in the crystals and those which are sometimes said to be in the "gaseous state".

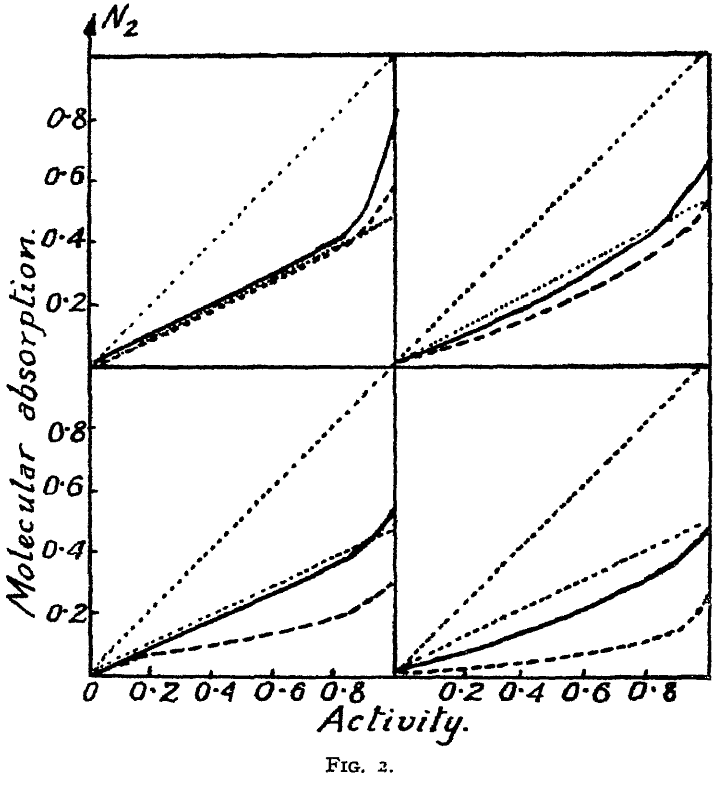


FIG. 2.

It is interesting here to compare our experiments on the reactions of solvents on the nitrocellulose with those of Champetier[12] on the fixation

---

[11] Hermans, *Kolloid Z.*, 1941, 96, 311.
[12] Champetier, *Ann. Chim.*, 1933, 20, 5.

of water on cellulose. Applying a purely chemical method derived from the method of residues of Schreinemader, Champetier finds that native cellulose immersed in water fixes one molecule of water for every pair of C_6 groups. These results can be interpreted as indicating the existence of a hydrate of cellulose—one cellobiose group to one molecule of water but, as Champetier as well as other workers have noted, the cellulose X-ray diagrams remain unaltered. No one believes that water can penetrate the crystalline lattice. If Champetier's results have rarely been made use of, the reason is probably that the correct interpretation of them has not yet been given. It should be noted that the method employed by Champetier is probably a precise means of estimating the extent of the zones of disorder and of crystalline regularity in any substance whenever a reagent adapted to the purpose is available.*

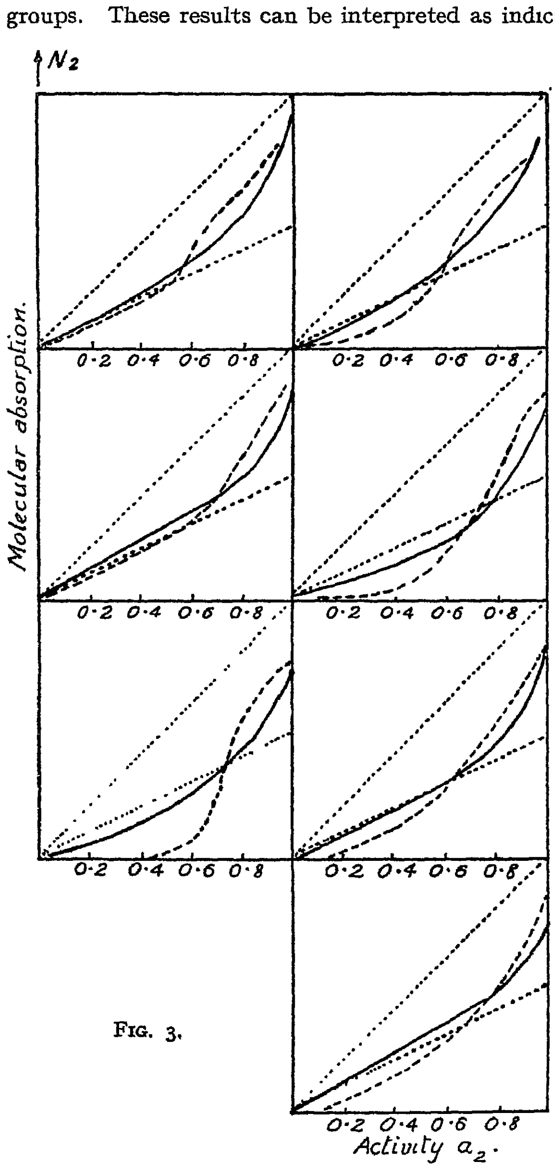


FIG. 3.

Intermediate Cases.

The Organic Esters.—The reactions of acetone and of methyl nitrate represent extreme types; one can readily conceive of intermediate cases. In one of these there is, to begin with, an absorption localised in the zones of disorder of the fibre structures, and then, for a certain activity of the solvent, absorption begins in the region of regularity. This is what we observe with solvents where the molecules are large enough to be impeded—propyl acetate, methyl butyrate, butyl acetate, methyl valerate, ethyl butyrate, propyl propionate, amyl acetate.

As for methylnitrate, the absorptions begin by following very closely the

---

* See the contribution of Champetier and co-workers at the symposium on High Polymers held in Strasbourg, *J. Chim. Physique* (in press).

straight line which characterises this solvent; when the absorption into the fibres reaches the value 0·5, *the absorption curves rise and tend towards the point* $a = 1$, $N_2 = 1$. *At the same time the diagram of the nitrocellulose remains unchanged up to the absorption of* 0·5 *and then changes abruptly above this concentration.* Another intermediate case is that in which *the absorption proceeds in both zones at the commencement but more readily in the disordered than in the ordered regions.* Here the absorption curve is situated between the two lines characteristic of acetone and of methyl nitrate respectively. Esters of organic acids with relatively small molecules such as methyl and ethyl acetate, and methyl propionate are examples of this.

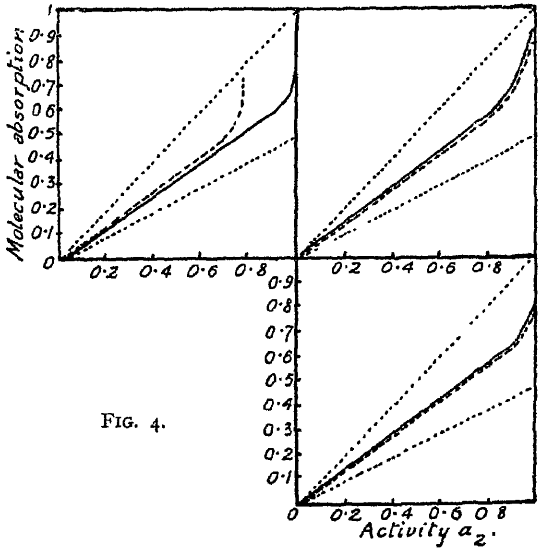


FIG. 4.

In these cases X-ray diagrams change at the beginning of the absorption, but not in such a definite way as for acetone. Often a mixture of phases can be observed—an unmodified nitrocellulose phase and a modified one. The latter appears clearly only when the concentration of the ester in the fibre exceeds 0·5.

Camphor.—The case of camphor is particularly remarkable. With camphor dissolved in an inert solvent such as hexane or petroleum ether, there is found to be a critical concentration below which the camphor is not absorbed by the nitrocellulose. At this concentration (about 40 % of camphor in the solvent) the absorption abruptly assumes a considerably higher value—about one molecule of camphor to one C_6 group.[13] In camphor vapour under the same conditions, absorption does not take place at all, even in the pure vapour of camphor. The molecules of camphor have difficulty, therefore, in penetrating into the nitrocellulose fibres. This difficulty is removed if the absorption occurs in the presence of methyl alcohol. It then takes place, even in the vapour phase. Fibres of trinitrocellulose were exposed to the mixed vapours of methyl alcohol and camphor at 40° or 50°. The result is an absorption curve which indicates the amount of camphor absorbed (N_2) as a function of the composition of the camphor-alcohol solution.

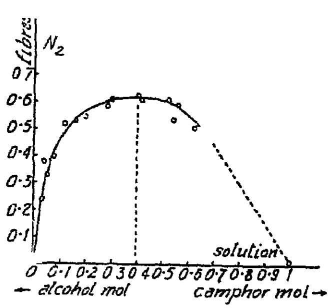


FIG. 5.

The absorption attains 0·6 molecule of camphor, and the X-ray diagram shows a change of structure. In presence of the methyl alcohol, the intimate structure of the fibres opens and the molecular chains separate to allow the camphor to penetrate the region of crystalline order. This change of structure does not appear until the proportion of one molecule of camphor to one C_6 group is exceeded. The intra-crystalline absorption does not begin, therefore, until all the zones of disorder have been saturated.

---

[13] Desmaroux, *Mémorial Poudres*, 1928, 23.

So with molecules such as those of camphor, absorption even into the disordered regions is not possible except in presence of small molecules which seem to open the road to the larger molecules. The substance with the smaller molecules therefore acts as a catalyst for the absorption reaction.*

II. Densities.

The determinations of density of cellulose derivatives in various liquids give information about the penetration of the liquids into the fibre-structure. Generally speaking from a knowledge of the structure and composition of a phase, its density can be calculated. This density is generally slightly higher than that measured by the usual methods (even if precautions are taken to remove the air-bubbles adherent to the solid) because the crystalline solid has a mosaic structure in which there are voids between the mosaic blocks, and into which the liquid cannot find its way.

If the measured density becomes identical with the calculated density, either the solid has not a mosaic structure, or the liquid used for the measurement penetrates into the empty spaces. Thus if the determinations of density in a liquid gave values systematically higher than the theoretical values, it must be considered that the liquid penetrates into the spaces in the structure of the regions of crystalline order. In the case of such a material as cellulose where there are regions of both order and disorder, and where the regions of order are themselves arranged into a very pronounced mosaic structure, special difficulties must be expected in the performance and interpretations of density measurements.

We have made determinations and shall now review the conclusion that may be drawn from them. We were unaware that similar investigations had been made in Holland by Hermans.[15] Our measurements are much less precise than those of the Dutch author, but they perhaps extend over a larger field and we shall therefore give them, and compare them in each case with those of Hermans.

Cellulose.—Accepting for the dimensions of the unit cell, the most recent values of Kurt Meyer: $a = 8{\cdot}35$, $b = 10{\cdot}3$, $c = 7{\cdot}9$, $\beta = 84°$, with $M = 162$ and $1{\cdot}66 \times 10^{-24}$ g. for the weight of the hydrogen atom, the density of cellulose is calculated to be 1·582. Making use of a flotation method with carbon tetrachloride after drying the fibres at 100° in a current of air, Hermans found the density of ramie to be 1·553, and that of bleached cotton to be 1·545.

Working with a simple pyknometer and using various liquids supposed to have no action on the cellulose, we have obtained the results in Table I The cellulose was bleached ramie dried *in vacuo* over sulphuric acid. The measurements were made at 20° c.

From these, one can deduce that organic liquids give a normal cellulose density. Hexane, however, should effect a deep penetration of the mosaic structure, entering between the domains of order. The almost theoretical value found with this liquid can be accounted for in this way.

When water is used, an evident penetration of the liquid into the structure of the fibres can be recognised. Hermans has made some measurements precise enough to show that the preliminary drying of cellulose fibres—ramie or cotton—has a marked influence on the results. Care must be taken not to dry the cellulose completely at the outset if the limiting density is to be found, but to take the precaution of leaving the fibres immersed in water for at least 16 hours, or of submitting them to a preliminary treatment with water vapour. If the fibres are dried at 100° before the measurement, the density found is always lower than the limiting value by 3 to 5 units in the third decimal place. The limiting values obtained by Hermans were 1·6116 for ramie and 1·6108 for standard cotton.

---

* The absorption of camphor has been studied by Petitpas.[14]

[14] Petitpas, *Mémorial Services Chim. de l'Etat.*, 1944, **31**, 260; *Compt. Rend.*, 1944, **218**, 785.

[15] Hermans, *Contribution to the Physics of Cellulose Fibres* (Elsevier Publishing Co., 1946).

We made our measurements on fibres dried over sulphuric acid *in vacuo* at room temperature. When the density determination was made at once, values of 1·590, 1·598 were found; after waiting 24 hr. to allow water to penetrate, *d* was 1·603 ; for cotton, when the measurement is made immediately, *d* = 1·597. *Our measurements are thus quite comparable with those of Hermans* and it can be affirmed that the measured value in water is slightly higher than the calculated one. *It must be concluded that the water penetrates between the cellulose chains.*

TABLE I.

DENSITY OF CELLULOSE IN ORGANIC LIQUIDS.

Liquid	d_m
Hexane . .	1·589 1·592
Benzene . .	1·555 1·567
Vaseline oil .	1·544

Suppose that there occurs a penetration of water into the regions of regularity in the proportion of half a molecule of water to each C_6 group in accord with an interpretation which may be given to the results of Champetier, the calculated density will be considerably higher than the measured one. The explanation of the slight difference observed is readily found in the penetration of water *into the vacant spaces of the mosaic structure and between the chains in the disordered regions.*

If the density is modified by heating, it constitutes a proof of the fact that elevation of temperature in the absence of water causes the chains to adhere to each other by aggregation of adjacent hydroxyl groups. These aggregations are of the same nature as those which exist in the regions of regularity and which maintain the cohesion of the chain in the lattice of the crystal. By multiplying the points of contact between neighbouring chains, the penetration of water between these chains is diminished. These deductions from the densities are in accordance with what is known of the so-called " water of swelling ".*

Trinitrocellulose.—Taking for the dimensions of the unit cell the values $a = 12{\cdot}5$, $b = 25{\cdot}6$, $c = 9$ and $\beta = 90°$,† the calculated density for trinitrocellulose of 14 % nitrogen, and molecular weight 294 for a C_6 group, is equal to 1·685.

The values obtained in different liquids at 20° by the classical method of the pyknometer are given in Table II.

The experimental values are smaller than the calculated, except for hexane and alcohol. In cellulose or in the nitrocellulose which is derived from it, hexane will penetrate in the same way into the *lacunae* of the mosaic structure.

---

* What has been called the " water of swelling " is that water which remains in the fibres after vigorous centrifuging. This water can be got rid of only by heating to 100°. It is known that the amount of this " water of swelling " in fibres of natural cellulose (cotton or ramie) and of viscose depends especially on the treatments the fibres have received.

† It should be noted that in a previous paper there was an error in calculation of the unit cell of trinitrocellulose. The X-ray diagrams should be referred to a lattice of which the dimensions are those given above. These values correct those which have been published: Mathieu, *Structure des Nitrocelluloses* (Herman Co., Paris, 1935), p. 97; *Memorial Services Chim.*, 1943 (Th. Petitpas), p. 252

In its relation to the nitrocelluloses, alcohol behaves as water does to cellulose. Remarks similar to this, have often been made in the technology of the cellulose industries; for instance, paper can be made either from cellulose with water, or from nitrocellulose with alcohol.

TABLE II.

DENSITIES OF NITROCELLULOSE OF 14 % NITROGEN.

Liquid.	Density of the Liquid at 20°	Density of Nitroramie.
Vaseline oil .	0·859	1·608
Benzene .	0·877	1·636
Cyclohexane	0·784	1·644
Water . .	0·9994	1·669
Hexane .	0·675	1·686
Ethyl alcohol	0·800	1·696

Calculated density $D_c = 1{\cdot}685$

Nitrocelluloses of Varying Nitrogen Content.—Finally we come to the density determinations of substances with *lacunae* in their structures. The nitrocelluloses are formed by the displacement of cellulose chains in such a way that —O . NO_2 groups can replace —OH groups. The displacement takes place first of all, and the nitrate groups progressively take their places in the new structure thus formed. Between the moment at which the chains are displaced and that at which the —OH groups are all esterified, the structure of the nitrocelluloses appear to be markedly lacunary. What would result, in this case, from the density measurements and the comparison of these measurements with those which can be calculated on the supposition that the structure is that described? Table III gives the answer to this question. The densities were measured at 20° c. From these results it can be concluded that water certainly enters the lacunary structure of the nitrocelluloses so long as these retain the structure of the dinitrocompound.

TABLE III.

DENSITIES IN WATER OF NITROCELLULOSES OF DIFFERENT NITROGEN CONTENT.

% Nitrogen.	$d_{\text{Calc. dinitro}}$ *	$d_{\text{Calc. trinitro}}$.	d_{measured}.
11·2	1·510		1·656–1·652
11·5	1·470		
11·9	1·452		1·659
12	1·450		
12·2	1·440		1·674
12·6	1·430		
13	1·40	1·705	
13·1		1·704	1·662
13·5		1·70	
13·7		1·695	
14		1·685	1·669

* The decrease in the calculated density arises from the variation of cell-dimension with nitrogen-content.

Phases Containing Nitrocellulose and Gelatiniser.—Fibres which have absorbed a solvent give density values which are particularly interesting. In this case the measurements must be made with fibres swollen with the solvent which they have absorbed, i.e. the measurement must be made under the very conditions of absorption, and the liquid which has been absorbed is therefore used as immersion liquid. The weight of the solid phase is found by testing the sum of the weights of the dry fibres and of the solvent fixed by the cellulose compound, the latter being deduced from the absorption curves already plotted. It is then sufficient to measure the volume of the fibres in the pyknometer at the temperature of experiment (40°).

TABLE IV

HEXANE—ACETONE

c.	d_1.	N_2.	d_c.
1·2	0·676	0·25	1·620
2·0	0·676	0·4	1·446
3·4	0·677	0·53	1·420
7·0	0·680	0·69	1·180
8·0	0·681	0·68	1·199
12·5	0·683	0·74	1·180
15	0·687	0·76	1·084
16·5	0·685	0·77	1·118
19	0·692	0·78	1·108
30	0·700	0·84	1·045
44	0·715	0·9	0·925
53	0·726	0·9	0·982
Hexane	0·673	0	1·686
Acetone	0·795	—	—

For the last measurement, so long as the concentration is not very high and the fibres not very much swollen, the absorption is carried out separately in small flasks and when the equilibrium is established, the fibres and the equilibrium liquid are transferred to the density flask. When the fibres are very swollen and form a gel, the absorption is carried out under the same conditions of temperature and duration, but in the density-vessel itself (the capillary and stopper of which are sealed to prevent evaporation) and the level is adjusted at the moment of weighing.

To state precisely how the measurement is made: let P be the weight of dry fibres, p, the weight of the solvent fixed and known from the absorption curves, d_1 the density of the liquid in equilibrium (measured separately) and P_1 the weight of liquid, then

$$d_e = \frac{P+p}{P_1} d_1.$$

This measurement of the density presupposes that the fibres are composed of cellulose derivatives and fixed gelatiniser only, and that the inert solvent will not be fixed at all in the fibres.

Tables IV and V refer to mixtures of acetone and hexane, and of cyclo-pentanone and hexane, where C is the concentration of acetone or cyclo-pentanone in 100 g. of active mixture, d_1 = density of mixture, N_2 = molecular concentration of ketone in the fibres, and d_e = density of the nitrocellulose ketone phase.

TABLE V.

CYCLOHEXANE-CYCLOPENTANONE.

c.	d_1.	N_2.	d_c.
4·0	0·788	0·32	1·607
6·0	0·791	0·40	1·550
11·0	0·798	0·55	1·476
21·0	0·812	0·65	1·370
38·0	0·838	—	—
100	0·952	0·985	0·970
Cyclohexane	0·784	0	1·644
Cyclopentanone	0·952	—	—

TABLE VI.

AFTER CALVERT.

N_2.	d_c.
0·36	1·53
0·63	1.30
0·675	1·24
0·755	1·16
0·80	1·10
0·82	1·05
0·86	1·05
0·95	0·9
1	0·81

Table VI gives the density values obtained by Calvert for homogeneous collodions made by dissolving dinitrocellulose in acetone; N_2 in that table denotes the molecular composition of the collodion.

It is evident from these tables that *the variation of the determined density, as we have defined it, with the value of N_2 is continuous from that of the dry fibres right up to that of the pure gelatiniser.* It is remarkable that the curves of density of the fibre phases are entirely in accord with those of the liquid collodions.

Let us now compare the measured densities, as described, with the densities calculated on the assumption of the structures described by Petitpas.[1] These structures, which become certain when $N_2 = 0{\cdot}5$, have their minimum of density for this value of N_2. It is for this concentration that the crystal structure is, in the highest degree, lacunary. As the *lacunae* are regularly filled with molecules of ketones, the density increases *pari passu*.

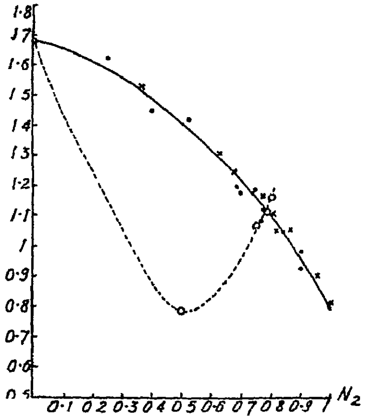


FIG. 6.

Fig. 6 shows the comparison of calculated and measured densities.

The comparison leads to the conclusion that there is a penetration *of the active solution into the interior of the lacunary structure.* The phase, where the composition is fixed by the absorption measurement, consists of a loose lattice in which the equilibrium liquid, and the mixture of ketone and inert liquid are dissolved. At the beginning of the reaction the *continuum* is the lattice of the cellulose material; when one gets to the collodion, *the continuum* is the liquid, the cellulose chains being dispersed in it.

It may be noted that in the two series of experiments the abscissa of the point at which *the calculated equals the measured density is precisely the value of N_2 at which the dispersion of the fibres is observed.*

Methyl Nitrate.—If the density of a trinitrocellulose of 14 % nitrogen is measured in pure methyl nitrate under similar conditions to that of cellulose in water, i.e. no absorption of the nitrate into the zones of regularity is supposed to occur, the density found is 1·870, which is much greater than the calculated one (1·685). This points to a remarkable fixation of the methyl nitrate not only in the lacunary spaces of the mosaic structure, but also on the parts of the chains which are in the region of disorder. If the density deduced from these values is interpreted as that of a phase into which enter two molecules of nitrate for each C_6 group, and if the calculation is made as it has already been made for acetone and cyclopentanone, $P + p$ being used as the weight of the solid phase, the value of the measured density which results is 1·540. Since the penetration of two molecules of nitrate into the crystalline structure takes place without altering the structure, the calculation leads to a value of 2·3. This hypothesis is very improbable, and we are therefore led to the confirmation of what has been already supposed—*methyl nitrate is fixed in the regions of disorder, on the surface of the crystalline micelles, and in the zones of irregular arrangements.*

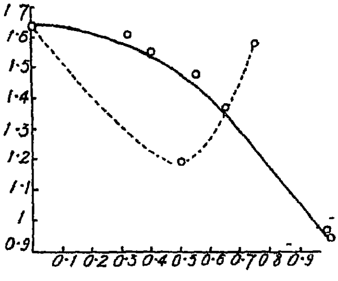


FIG. 7.

Summary and Conclusion.

In conclusion we may state that studies of absorption of various reactive substances by a fibrous system, examinations of structure by X-rays, and determinations of densities, lead to certain conclusions relating to the penetration of the structure of fibres by the reacting substances.

It follows that the action of such a substance can be localised. Thus we can distinguish three types of gelatiniser for cellulose derivatives.

(1) Those which attack the molecular chains and whose action is not affected by the cohesion of these chains in the arrangements which are based on an ordered crystalline lattice.

(2) Those which cannot undo the regular arrangement of the crystalline regions but which attack the free portions of the chains, the surfaces of the micelles and the regions of disordered arrangements.

(3) Finally those which are incapable alone, but whose action is " catalysed " by the presence of smaller molecules.

There exist all kinds of intermediate cases.

To these observations it would be interesting to add a number of observations made in connection with the manufactures of powders for rifles and guns and of paper-making.

These researches have been made under the direction of M. l'Ingenieur General Jean Desmaroux. They have been published with the permission of the Direction des Industries Chimiques du Ministere de la Production Industrielle.

Laboratoire Central des Services chimiques
de l'Etat.

GENERAL DISCUSSION *

Dr. R. J. B. Marsden (*Manchester*) (*communicated*) : In answer to Professor Bernal's point, as to what is the underlying pattern embracing all swelling systems, surely we already have it in the three criteria following, which a system must satisfy, for swelling to occur.

(i) There must be large units which can separate from each other while remaining definite entities, themselves not undergoing any breakdown. These large units may either be aggregates of small molecules as for instance in soaps; large molecules, either linear as in thermoplastic polymers, or laminar as in graphitic-type compounds; or even small crystals.

(ii) There must be agencies causing *localised* cohesion between these units, which may vary from true covalent cross links, the " crystalline " regions of cellulose and similar linear polymers, or the " near-ice " layers of small solvent molecules mentioned by Bernal. (Otherwise solution and not swelling will occur directly.)

(iii) A solvent which can release all the other forces which hold the large units together.

To take a specific example, it is probable that a diamond is unswellable, as if it were swollen it would no longer be a diamond, whereas graphite and graphitic bodies can swell, as in mellitic acid and the coals and cokes.

Further than this it does not seem reasonable to me to expect a detailed and universal explanation to go, if it is to cover as diverse systems as the ionic (Procter-Wilson) swelling of gelatin, and the swelling of polystyrene by benzene.[1]

Dr. J. J. Hermans (*Utrecht*) said : In connection with Dr. Hartley's work, I wish to draw attention to a diffusion phenomenon with very sharp boundary line resulting from a reaction of the diffusing substance with the medium into which it diffuses. Recent experiments, which are still in progress, have shown that a diffusion with sharp boundary line always occurs if the diffusing molecules are precipitated by the medium and thus withdrawn from the diffusion process. For example, if copper ions diffuse into a cellulose xanthate gel, the copper is fixed by xanthate groups. One observes a very sharp boundary, dark-brown in colour, which moves into the gel at a rate diminishing with time. Similar experiments have been carried out with $Na_2S_2O_3$ diffusing into a gel containing iodine. A very beautiful sharp boundary line is observed if Na_2S diffuses into a gel which contains lead acetate. In this case both the S and the Pb take part in the diffusing process. At the boundary, where these ions meet, PbS precipitates and we have the situation represented in the figure. The theory of this phenomenon predicts that the boundary line conforms to the simple relation $X/\sqrt{t} =$ const., this constant being a function of the diffusion coefficients of S and Pb and of the ratio a/b. This theory is completely borne out by experiment, and it would seem that this phenomenon is quite general in all cases where the diffusing molecules are caught by the medium and thus withdrawn from the diffusion process.

---

* On three preceding papers.

[1] Compare introduction to " The Swelling of Cellulose Acetate in Aqueous Phenol Solutions," Marsden and Urquhart, *Shirley Institute Memoirs*, 1941-42, **18**, 121, or *J. Text. Inst.*, 1942, **33**, T 105

It might be of interest to consider whether this phenomenon plays a part in some of the examples studied by Hartley, since it is well known that very often a real molecular compound is formed between a polymer and a solvent. In which case, of course, the discontinuity would be a real and essential feature of the process and not " an ordinary decreasingly sharp initial boundary between the two substances."

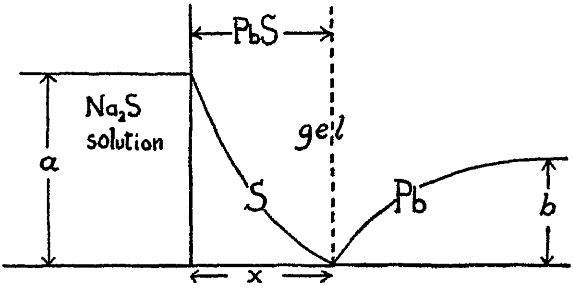


FIG. 1.

Dr. G. S. Hartley (*Maidenhead*) said: In reply to Dr. J. J. Hermans, he realised that the abnormal sharp form of the advancing diffusion front could have some specific mechanism in special cases such as those to which Dr. Hermans had referred. The conclusion that the phenomenon requires a diffusion coefficient falling rapidly with decreasing concentration of the advancing component is, however, a quite general one and does not rest on any particular mechanism. The mechanism which the author had in mind was the simple one of rapidly increasing viscosity and, in as much as the phenomenon is a very general one in polymer solvent systems, it appears preferable in the first place to regard it as the normal behaviour in such systems rather than to invoke adsorption or solvation effects which must be specific.

In reply to Dr. Dervichian, the expanding scale remained coherent up to an extension of at least three-fold and would probably extend further in the absence of convection, which was not avoided in the experiment shown on the screen.

Mr. J. Crank (*Maidenhead*) said: There is a reference in Dr. Hartley's paper to the familiar fact that the last stages of desorption frequently proceed more slowly than the last stages of absorption in the same system. It is clear that this must be so if the diffusion coefficient is variable and is greater at high concentrations of the diffusing substance than at low concentrations. It is not so easy, however, to see how the early stages of absorption and desorption depend on the relation between diffusion coefficient and concentration, since the formal mathematical treatment of the diffusion equation becomes much more difficult when a variable diffusion coefficient is involved. It is possible, however, for any given relation between diffusion coefficient d and concentration c, to evaluate approximate numerical solutions of the diffusion equation, using step-by-step methods of integration. In this way a number of complete absorption-time and desorption-time curves, which show one or two features of general interest, have been obtained. The actual calculations refer to diffusion in one dimension into a parallel-faced slab of finite thickness. The surface concentrations of diffusing substance have equal and constant finite values throughout absorption, and both are zero throughout desorption.

The figure shows the general behaviour of the absorption and desorption rate curves for the three types of variable diffusion coefficient indicated in the inset diagrams, where M is the amount of diffusing substance absorbed

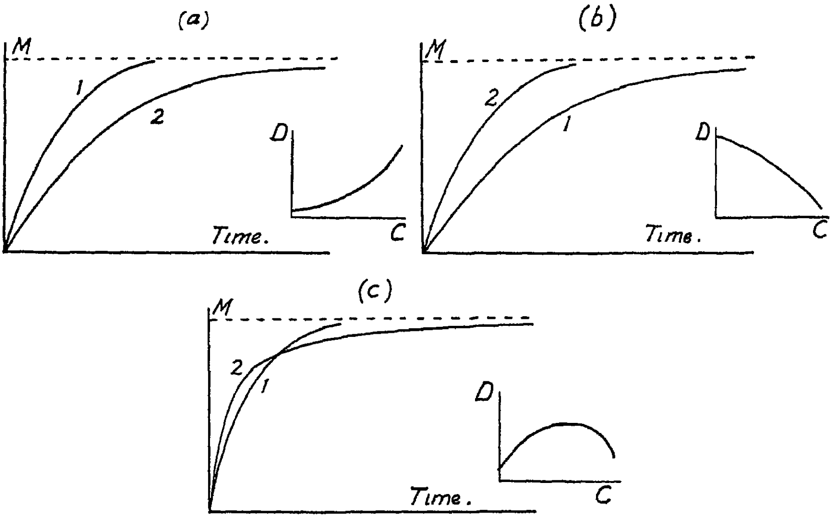


FIG. 1.—In each case curve 1 refers to absorption and curve 2 to desorption

or desorbed as the case may be. When D increases with C increasing throughout the relevant range in C (Fig. 1*a*), the desorption process is everywhere slower than the absorption process, whereas Fig. 1*b* shows the opposite to be true when D decreases as C increases throughout the range. Fig. 1*c* shows that the presence of a maximum in the curve relating D and C may cause the absorption and desorption curves to intersect. Desorption proceeds more rapidly than absorption in the early stages, but the final stages of desorption are again slower than those of absorption, provided the value of D at the highest value of C is greater than that at $C = 0$. Absorption and desorption rate curves of the types shown in Fig. 1*a*, 1*b* and 1*c* have been found experimentally when water is sorbed and desorbed by leather, rubber and horn respectively.

It is hoped later to publish quantitative solutions for a number of specific numerical relations between diffusion coefficient and concentration.

Dr. J. R. Scott (*Croydon*) (*communicated*) : An interesting example of a more or less sharp advancing boundary in the swelling of a high polymer (in this case an insoluble one) is found with ebonite, which can be regarded as a 3-dimensional network of polyisoprene chains with cross-links formed by sulphur. It has been shown by Scott [2] that with liquids that have a strong swelling action, e.g. benzene and its homologues, chlorobenzene, carbon disulphide, carbon tetrachloride, and nitrobenzene, the rate of liquid absorption soon becomes practically constant and remains so until the ebonite is saturated, when absorption ceases quite suddenly. This uniform rate of absorption clearly represents the steady inward advance of the boundary between unswollen and swollen solid. The papers referred to [1] gave, as the explanation of this, the same effect as is described in the more recent papers quoted by Hartley, namely, that the diffusion constant of the liquid through the solid is greatly increased when the latter is in the swollen condition.

Dr. P. H. Hermans (*Utrecht*) said : If a solvent-polymer system showing merely limited but not infinite swelling and final solution were selected for the experiments, the conditions would be simpler. The double refraction arising from unidirectional swelling (which is equivalent to that arising from unidirectional stretching) would also persist and not diminish by rotational diffusion. Quantitative comparison with double refraction caused by deformation would then be easier.

The phenomena described by Dr. Robinson are but another illustration of the intrinsic equivalence of swelling and deformation processes arising from the concept of molecular network structures. If a dry isotropic polymer is allowed to swell isotropically, the coiled chains of the network will be stretched without, however, giving rise to double refraction, no more than isotropic three-dimensional extension would. If isotropic swelling is impeded, as in Dr. Robinson's experiments, and only unidirectional swelling is allowed, the stretching of the chains takes place in one preferred direction and gives rise to optical anisotropy.

It should be remarked that, when quantitatively comparing the effect of unidirectional swelling with that of mechanical extension, as attempted by Dr. Robinson, the correct volume relations should be taken into account. If swollen gels are stretched, their volume changes in the majority of cases, either increasing (like in Kratky's experiments with swollen cellulose acetate) or decreasing (as with cellulose swollen in water). The displacements measured in Dr. Robinson's method are equivalent to extensions at constant volume. If the system shows volume increase on extension the equivalent at constant volume will be larger. This may be a possible explanation of the difference found with Kratky's results.

Dr. Conmar Robinson (*Maidenhead*) said : Some experiments have been carried out with methyl ethyl ketone diffusing into cellulose acetate, which swells, but does not dissolve, in this liquid. The double refraction curve obtained was qualitatively the same as with acetone, the peak of the curve remaining unchanged as it traverses the acetate, and relaxation taking place rapidly behind the peak ; here again, rotational diffusion diminishes the double refraction as soon as swelling has sufficiently diluted the polymer. For the purpose of the experiments there is no essential difference between the solvent on the one hand and the liquid which produces a limited degree of swelling on the other, except that in the latter case the outward moving boundary persists and will not eventually become invisible or be destroyed by convection.

The fact that relaxation takes place rapidly in the swelling gel as well as in the solution shows that to consider swelling as the deformation of a molecular network is an over-simplification. The polymer gel or solution is a molecular felt, any molecule of which has scores of points of contact with other polymer molecules, and the system is unlike a network in that these contacts are continually making and breaking. The degree and nature of these adhesions may suffice in preventing the polymer molecules escaping into solution (as in methyl ethyl ketone) while not preventing a return to random orientation following the deformation which accompanies swelling.

---

[2] Scott, *Trans. Inst. Rubber Ind.*, 1937, **13**, 109. Additional details, *J. Rubber Res.*, 1941, **10**, 123.

It was naturally of interest to make a comparison with Kratky's data, but it is best not to stress any apparent discrepancy until the preliminary results we have presented here have been replaced by the more accurate values which we hope to obtain.

THE INTERACTION BETWEEN RUBBER AND LIQUIDS.

X. SOME NEW EXPERIMENTAL TESTS OF A STATISTICAL THERMODYNAMIC THEORY OF RUBBER-LIQUID SYSTEMS.

BY GEOFFREY GEE.

Received 19*th June*, 1946.

In recent years a number of papers have been published, in which the thermodynamic properties of rubber liquid systems have been deduced by statistical methods. These theoretical treatments have on the whole been highly successful, but their critical evaluation is rendered difficult by the almost complete lack of reliable experimental data. Any reasonably comprehensive attempt to fill this gap would be extremely laborious. By concentrating attention on points where, if at all, weaknesses were likely to be found, it has been possible to review fairly quickly the present position.

Calculations of the Gibbs free energy of dilution ΔG_0 of a polymer liquid mixture by the liquid have been based on a lattice model of the system. A lattice point is assumed capable of accommodating either a molecule of solvent or a "segment" of the polymer chain equal in volume to a solvent molecule. An approximate treatment [1, 2] leads to a result of the form

$$\Delta G_0 = RT\{\ln (1 - v_r) + v_r + \mu v_r^2\} \quad . \quad . \quad . \quad (1)$$

where v_r is the volume fraction of rubber in the mixture, and μ is a parameter given by

$$\mu = A + B/RT \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where A and B are constants, and Bv_r^2 is the heat of dilution, ΔH_0. The only extensive data at present available to test this result are those for the system rubber and benzene.[3, 4] The results at 25° C. are represented by equation (1) with only small variations of μ over the whole range of composition, but A and B are both strongly dependent on v_r, especially when this is low.[4] Huggins [5] has shown that equation (1) describes the rather fragmentary data for a number of rubber-liquid systems at a single temperature. As a result there is a growing tendency to employ μ as a quantity characterising a particular polymer liquid mixture. The simplicity of such a procedure makes it very attractive, but it is clearly desirable to apply some critical experimental tests of its validity. This is the purpose of the present paper and an experimental study has therefore been made of some two-phase equilibria.

---

[1] Flory, *J. Chem. Physics*, 1942, **10**, 51.

[2] Huggins, *Ann. N.Y. Acad. Sci.*, 1942, **43**, 1.

[3] Gee and Treloar, *Trans. Faraday Soc.*, 1942, **38**, 147.

[4] Gee and Orr, *ibid.* (in press).

[5] Huggins, *Ind. Eng. Chem.*, 1943, **35**, 216.

The experimental work to be described is largely concerned with vulcanised natural rubber. It is assumed that the introduction of a few cross links into the rubber will have a negligible effect on the heat and entropy of mixing. There will, however, be an additional term to be added to equation (1) representing the decrease of entropy resulting from the network deformation consequent upon imbibition of liquid. The correction has been evaluated by Flory and Rehner [6] in terms of the molecular weight M_c between junction points of the network. For isotropic swelling they obtain

$$\Delta G_0 \quad RT\left\{\ln(1 - v_r) + v_r + \mu v_r^2 + \frac{\rho_r V_0}{M_c} . v_r^{\frac{1}{3}}\right\} \quad . \quad . \quad (3)$$

where ρ_r is the density of the rubber, and V_0 the molar volume of the liquid. If a piece of rubber is constrained to a length l, compared with an initial length l_0 dry and unstrained, equation (3) must be replaced by

$$\Delta G_0 = RT\{\ln(1 - v_r) + v_r + \mu v_r^2 + \rho_r V_0 l_0/lM_c\} \quad . \quad . \quad (4)$$

The elastic tension f in the swollen rubber is given by [6, 7]

$$f = \rho_r A_0 RT(l/l_0 - l_0^2/l^2 v_r)/M_c \quad . \quad . \quad . \quad (5)$$

where A_0 is the cross-section of the dry unstrained rubber. Equation (5) has been shown [8] to hold for highly swollen rubbers, although it is less satisfactory when v_r is large. If, therefore, M_c is estimated from the tension in a highly swollen rubber, there is good reason to expect the last term of equations (3) and (4) to be accurate for small v_r. Since this term is only important when v_r is small, it is considered that data on vulcanised rubbers may safely be used to test equation (1).

The quantity measured in the experiments reported below is the equilibrium volume of liquid Q_m imbibed by unit volume of the rubber, and three types of measurement are included. These give the dependence of Q_m on the degree of cross linking, on the degree of extension l/l_0 of the rubber, and on the temperature. Equations (3) and (4) readily lead to the following expressions for these quantities when the equilibrium condition $\Delta G_0 = 0$ is inserted.

For isotropic swelling, the effect of M_c is given by

$$\ln(1 + 1/Q_m) - (Q_m + 1)^{-1} - \mu(Q_m + 1)^{-2} = \rho_r V_0/M_c(Q_m + 1)^{\frac{1}{3}}. \quad (6)$$

The effect of elongation is given by

$$\ln(1 + 1/Q_m) - (Q_m + 1)^{-1} - \mu(Q_m + 1)^{-2} = \rho_r V_0 l_0/lM_c. \quad . \quad (7)$$

Reference will also be made to some earlier data on the dependence of the tension f on the degree of swelling $Q(= 1/v_r - 1)$; an expression for this is obtained from equation (5) :

$$(\partial f/\partial Q)_l = \rho_r A_0 RT l_0^2/l^3 M_c. \quad . \quad . \quad . \quad . \quad (8)$$

In discussing the experimental data, use will be made of the following thermodynamic relations, derivations of which are given in Appendix I.

$$l_0\left(\frac{\partial Q_m}{\partial l}\right)_T = -\frac{V_0}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l \Big/ \left(\frac{\partial \Delta G_0}{\partial Q}\right)_l \quad . \quad . \quad . \quad . \quad (9)$$

$$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_{T, f = 0} = -\frac{V_0}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l \left\{\frac{1}{l_0 \partial Q_m/\partial l} - \tfrac{1}{3}(Q_m + 1)^{-\frac{4}{3}}\right\} \quad . \quad (10)$$

$$\Delta H_0 = T\Delta S_0 = -\frac{Q_m}{T}\left(\frac{\partial \Delta G_0}{\partial Q}\right)_T . \frac{\partial \ln(\Delta w)}{\partial\, 1/T} . \quad . \quad . \quad . \quad (11)$$

All the above are true only at $Q = Q_m$. In equation (11) Δw is the weight of liquid imbibed by a fixed amount of rubber.

---

[6] Flory and Rehner, *J. Chem. Physics*, 1943, 11, 521.

[7] James and Guth, *ibid.*, 1943, 11, 455.

[8] Gee, *Trans. Faraday Soc.* (in press).

Equations (9) and (11) show that the dependence of the maximum swelling on either length or temperature is inversely proportional to $\partial \Delta G_0/\partial Q$. Although the dependence of Q_m on cross linking is not strictly a thermodynamic problem at all, a simple analysis leads to the conclusion that this also is inversely proportional to $\partial \Delta G_0/\partial Q$. In general, therefore, the variation of Q_m with internal or external parameters is determined by the *slope* of the free energy-composition curve in the neighbourhood of the two phase equilibrium. It is this fact which makes the ability of an equation to describe these phenomena such a severe test of the theory on which the equation is based.

Effect of Cross-linking on Swelling Capacity.—Several natural rubber vulcanisates were prepared; the compounds, whose compositions are given in Appendix II, being selected to cover a range of "cure." The estimation of M_c was made by allowing a strip of rubber, carrying a dead load, to swell to its maximum capacity in 40/60 petrol ether at 25° C.

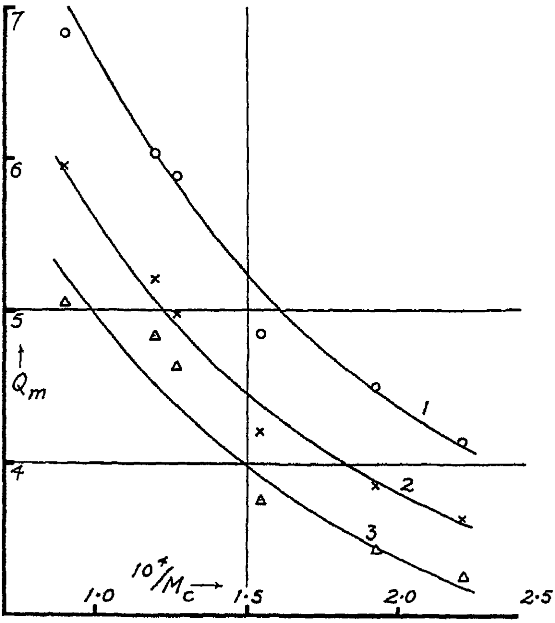


FIG. 1.—Effect of vulcanisation on swelling.
(*a*) Good swelling agents. 1. Carbon tetrachloride.
2. Carbon disulphide. 3. Benzene.
(Curves theoretical.)

When equilibrium was attained, the distance between two reference points on the strip was measured, and the degree of swelling estimated by cutting out the portion of the strip between the reference points and determining its petrol content. The value of $10^{-3} M_c$ obtained in this way are given in Appendix III; their reciprocals will be seen to run roughly parallel with the loads at 300 % estension as determined on a standard tensile testing machine (the so-called "modulus at 300 %"). Swelling measurements were carried out in a variety of liquids differing widely both in chemical type and in solvent power. Weighed samples of rubber (*ca.* 1 g.) were immersed in excess liquid and maintained at constant temperature for three days, by which time swelling was substantially complete, and then quickly surface dried and weighed. The results are shown in Fig.

1 and 2, in which Q_m is plotted as a function of $1/M_c$ for the various rubbers, with the addition, for the non-solvents, of unvulcanised rubber. Data for the two pairs of liquids, chloroform and carbon tetrachloride, benzene and toluene, were so similar that only one of each pair is shown. In order to fit the data to equation (6) a value of μ has to be selected to give the best general representation of the points, it being assumed that chemical differences among the rubbers would not be sufficient to affect μ significantly. The curves drawn in Fig. 1 and 2 were calculated from equation (6) by using the values of μ shown in Table I, and evidently give a fairly good representation of the data for the good swelling agents, bearing in mind the difficulty of making satisfactory comparisons of different rubbers. This is in agreement with the conclusion drawn by Flory[9] from his data on the swelling of a range of butyl rubbers in cyclohexane. The data for methyl ethyl ketone and ethyl acetate show a systematic deviation from the calculated curve, the effect of cross linking in reducing the swelling power being more marked than would be expected. It is difficult to estimate the quantitative discrepancy in δQ_m, the amount by which Q_m is lowered by cross linking, but it would seem that the experimental value for ethyl acetate is at least twice the calculated. It appears rather unlikely that any large part of this factor can arise from the term $\rho_r V_0/M_c(Q_m + 1)^{\frac{1}{3}}$, since it has previously been found that rubbers swollen by 100 % show nearly the theoretical behaviour in tension.[8] The tentative conclusion would therefore be that $\partial \Delta G_0/\partial Q$ for poor swelling agents is lower than would be expected on the basis of their swelling capacity; in other words, that equation (1) is only approximately true. The discrepancy involves only small deviations in the value of μ given by equation (6) when applied to different rubbers. Thus, e.g. Table I gives $\mu = 0{\cdot}78$ for ethyl acetate; the value calculated from the swelling capacity of raw rubber is $\mu = 0{\cdot}71$. Table I includes for comparison μ values calculated by Huggins[5] from vapour pressure data on raw rubbers, and the general order of agreement is fairly good, suggesting

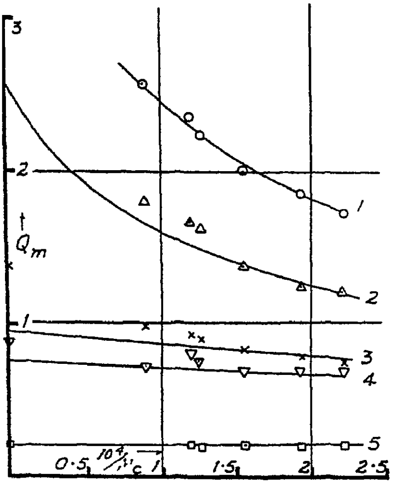


FIG. 2.—Effect of vulcanisation on swelling.

(*b*) Moderate and poor swelling agents.
1. 40/60 petrol ether. 2. *n*-Propyl acetate. 3. Ethyl acetate. 4. Methyl ethyl ketone. 5. Acetone.
(Curves theoretical.)

TABLE I.—VALUES OF μ FROM EQUATION (6).

Liquid.	μ (from Q_m).	μ (Vap. Press.).
Carbon tetrachloride .	0·29	0·28
Chloroform . . .	0·34	0·37
Carbon disulphide . .	0·425	0·49
Benzene	0·395	0·43 *
Toluene	0·36	0·43-0·44
40/60 petrol ether . .	0·54	0·43 †
n-Propyl acetate . .	0·62	—
Ethyl acetate . . .	0·78	—
Methyl ethyl ketone .	0·94	—
Acetone	1·37	—

---

* Data of Gee and Orr give 0·41. † Light petroleum.

[9] Flory, *Chem. Rev.*, 1944, 35, 51.

that for good swelling agents equations (1) and (6) are probably both reasonably satisfactory.

Effect of Extension on Swelling.—The measurement of the effect of extension on swelling is in principle much more satisfactory than the comparison of different rubbers. In practice it turns out to be very difficult except for liquids of moderate swelling power. With good swelling agents, only very small extensions can be applied without breaking the specimen, and it becomes necessary to measure relatively small differences in Q_m. In some cases, notably chloroform and toluene, the difficulty is enhanced by failure to attain steady equilibrium values of Q_m, the result obtained increasing very markedly with time.* With acetone the effect of extension on Q_m was very small, but with intermediate liquids, such as ethyl acetate and methyl ethyl ketone, very satisfactory and reproducible results were obtained over a wide range of extensions. Most of the data reported were obtained by stretching rings over glass formers, but for the better swelling agents, there was found to be less danger of rupture if the rubber was stretched in the form of a strip. This latter procedure was more troublesome, as it was necessary to cut out the stretched portion when swollen to equilibrium and determine its swollen weight. The sample was then returned to the liquid to determine its weight when swollen and unstrained, while finally the dry weight was obtained by removing the liquid *in vacuo*. By this method $\partial Q_m/\partial l$ is certainly underestimated for liquids in which Q_m increases with time.

The results are shown in Fig. 3 and 4, in which the theoretical curves are calculated from equation (7), using values of μ chosen to give the observed swelling at $f = 0$. Bearing in mind the above note about chloroform and toluene, equation (7) is satisfactorily confirmed for good swelling agents (Fig. 3). For poor swelling agents, the effect of extension is far larger than predicted. Flory and Rehner [11] have shown that equation (7) describes accurately the effect of extension on the swelling of a sample of butyl rubber in xylene.

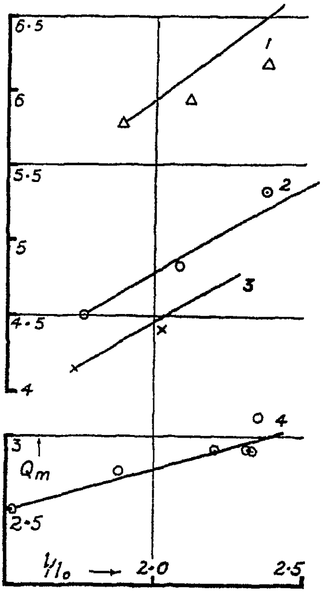


FIG. 3.—Effect of extension on swelling.
(*a*) Good swelling agents.
1. Chloroform. 2. Benzene.
3. Toluene. 4. 40/60 pet. ether.
(Curves theoretical.)

It is possible to determine whether the unexpectedly large dependence of Q_m on extension arises mainly from anomalies in the elastic behaviour or in the state of mixing, by making use of equation (9), but to do so it is necessary to estimate $(\partial f/\partial Q)_l$. Data have been reported previously [8] on the effect of toluene on the tension of stretched rubber strips; from these $(\partial f/\partial Q)_l$ is readily obtained by graphical interpolation. The resulting figures have been applied to the present data, this being justified by two facts: (*a*) that similar samples of rubber were used in the two investigations; and (*b*) that unpublished work † has shown that the tension at constant length of swollen rubbers depends, to a first approximation, only on the degree of swelling, and not on the nature of the liquid. It must be emphasised that the figures will not necessarily be very precise, since a considerable amount of graphical interpolation has been necessary. Some typical results are given in Table II, in which

---

* Scott [10] has reported that this is a normal feature of swelling, and has termed this relatively slow rise of Q_m the "increment". By working in nitrogen, and restricting the duration of experiments, no trouble has been encountered with most of the liquids used.

† Both by the author and by Dr. L. R. G. Treloar.

[10] Scott, *Trans. Inst. Rubber Ind.*, 1929, **5**, 95.

[11] Flory and Rehner, *J. Chem. Physics*, 1944, **12**, 412.

$$l_0\frac{\partial Q_m}{\partial l},\quad -\frac{1}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l \quad\text{and}\quad \left(\frac{\partial \Delta G_0}{\partial Q}\right)_l$$

are compared with values computed from the above statistical equations.

In considering these results, it must be borne in mind that $\frac{\partial Q_m}{\partial l}$ and $\left(\frac{\partial f}{\partial Q}\right)_l$ are both differential quantities, not measured directly, so that both are liable to considerable experimental error. None the less, it is clear from the table, that the high values of $\partial Q_m/\partial l$ arise mainly from very low values of $\left(\frac{\partial \Delta G_0}{\partial Q}\right)_l$. An important observation is that these low values are not essentially dependent on extension, the discrepancy between theory and experiment being of the same order at different extensions. This conclusion is confirmed by extrapolating the data to $f = 0$, and employing equation (10) to calculate $\partial \Delta G_0/\partial Q$ for the unstretched rubber; the results are set out in Table III. The figures for methyl ethyl ketone and ethyl acetate are consistent with the observed relatively large effect of vulcanisation on the swelling in these two liquids (Fig. 2).

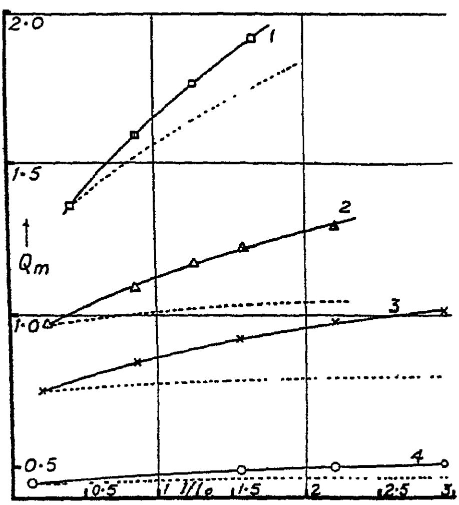


FIG. 4.—Effect of extension on maximum swelling.

(*b*) Poor swelling agents.

1. *n*-Propyl acetate (add 0·5 to ordinates).
2. Ethyl acetate. 3. Methyl ethyl ketone.
4. Ethyl formate.

(Dotted curves theoretical.)

By considering the magnitudes of the separate terms in equation (3) it is easily seen that $(\partial \Delta G_0/\partial Q)_{f=0}$ is determined essentially by the free energy of mixing, no important contribution arising from network deformation. It follows that the last column of Table III gives an estimate of the ratio between the true and calculated values of $(\partial \Delta G_0/\partial Q)_T$ for raw rubber at the values of Q shown in column 2. It is known by direct measurement that this ratio is close to unity for benzene,[3, 4] at all values of Q. The success of the statistical equations in describing both the elastic behaviour and the effect of extension on swelling for a series of good swelling agents shows that the same must be true of these. In contrast, the results of Table III suggest that as the limit of imbibition is approached for poor swelling agents, $\partial \Delta G_0/\partial Q$ falls below its calculated value. This has already been shown to be the case for methyl alcohol[12] where, in the neighbourhood of maximum swelling, $\partial \Delta G_0/\partial Q$ falls to not more than one-twentieth of the value to be expected from equation (4) and the

[12] Ferry, Gee and Treloar, *Trans. Faraday Soc.*, 1945, 41, 340.

observed maximum swelling. Lens' data for acetone [13] are not sufficiently precise to estimate the ratio (observed/calculated) with any precision, but it is certainly low ($\approx 0{\cdot}05$). These various observations are collected

TABLE II.—RELATION BETWEEN SWELLING, TENSION AND EXTENSION. (RUBBER A.)

Liquid.	Q_m (unstrained).	l/l_0.	$l_0 \frac{\partial Q_m}{\partial l}$.		$-\frac{1}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l$. Cal./cc. Rubber.		$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_l$. Cal./mole Liquid.	
			Expt.	Theory.	Expt.	Theory.	Expt.	Theory.
Ethyl formate	0·447	1·5	0·036	0·0105	$0{\cdot}052_5$	0·0310	11_7	237
		2·0	0·025	0·0061	$0{\cdot}036_5$	0·0175	11_7	230
		2·5	$0{\cdot}019_8$	0·0040	$0{\cdot}027_0$	0·0112	11_1	227
Methyl ethyl ketone	0·75	1·5	0·144	0·0348	$0{\cdot}043_2$	0·0310	27	80·1
		2·0	0·119	0·0218	$0{\cdot}025_9$	0·0175	22	75·8
		2·5	0·095	0·0138	$0{\cdot}017_6$	0·0112	17	73·1
Ethyl acetate	0·96	1·5	$0{\cdot}22_5$	0·074	$0{\cdot}039_7$	0·0310	17	40·1
		2·0	$0{\cdot}18_7$	$0{\cdot}047_5$	$0{\cdot}022_8$	0·0175	12	36·2
		2·5	$0{\cdot}15_0$	$0{\cdot}032_5$	$0{\cdot}015_0$	0·0112	10	33·7
n-Propyl acetate	1·86	1·5	0·54	0·40	$0{\cdot}033_5$	0·0310	$7{\cdot}_2$	$8{\cdot}9_3$
		2·0	0·45	$0{\cdot}30_3$	0·018	0·0175	$4{\cdot}_8$	$6{\cdot}6_7$
		2·5	0·36	$0{\cdot}24_8$	$0{\cdot}011_5$	0·0112	$3{\cdot}_7$	$5{\cdot}2_4$

in Fig. 5, in which the ratio of observation to calculation at the point of maximum swelling is plotted as a function of Q_m. There is no reason to expect all the points to lie on a smooth curve, but it seems extremely probable that the curve drawn will furnish at least a rough estimate of the ratio for other liquids. *There is, in other words, strong evidence that equation* (1) *fails to give the correct dependence of the free energy of dilution on composition for liquids of limited swelling power.*

TABLE III.—EVALUATION OF $(\partial \Delta G_0/\partial Q)_{f=0}$ (RUBBER A).

Liquid.	Q_m.	l/l_0 $(=(Q_m+1)^{\frac{1}{3}})$.	$-\frac{1}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l$. Cal./cc. Rubber.	$l_0 \frac{\partial Q_m}{\partial l}$.	$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_{f=0}$. Cal./mole Liquid.		Ratio Expt./Theory.
					Expt.	Theory.	
Ethyl formate .	0·447	1·13	0·075	$0{\cdot}04_9$	120	242	0·50
Methyl ethyl ketone	0·75	1·20	$0{\cdot}062_5$	$0{\cdot}15_9$	34	83·9	0·41
Ethyl acetate .	0·96	1·25	0·055	$0{\cdot}24_5$	21	42·2	0·50
n-Propyl acetate .	1·26	1·42	$0{\cdot}037_5$	0·55	$7{\cdot}_2$	$8{\cdot}1_5$	0·88

The way in which this failure accounts for some of the anomalous effects of vulcanisation and extension on swelling is shown schematically in Fig. 6, which relates to a liquid in which raw rubber is swollen to $Q_m = 1{\cdot}00$. Curves 1 and 2 represent the "experimental" and calculated curves of ΔG_0 while 3 and 4 give (minus) the increment due to vulcanisation, unstretched and stretched to $l/l_0 = 2$. The points of intersection of curves 3 and 4 with 1 and 2 give for the maximum swelling of the vulcanised rubber:

---

[13] Lens, *Rec. trav. chim.*, 1932, 51, 971.

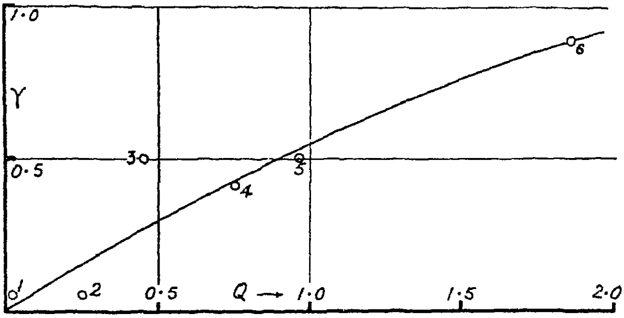


FIG. 5.—Ratio (γ) of experimental to theoretical values of $(\partial \Delta G_0/\partial Q)_T$.
1. Methyl alcohol. 2. Acetone.

unstretched,	experiment	0·828,	calculated	0·902
stretched,	experiment	0·876	calculated	0·930

(These figures are, of course, only illustrative, and have no quantitative significance.)

The Effect of Temperature on Swelling.—The temperature dependence of the maximum swelling of several natural rubber vulcanisates was determined in a range of liquids, employing the technique already described. In most cases measurements were carried out in duplicate, the first sample being allowed to swell successively at 0° C., 25° C., 0° C.; and the second sample at 25° C., 0° C., 25° C. In this way an adequate check was obtained on the attainment of equilibrium, and the absence of any important drift of Q_m with time. The results may be expressed in the form $-\partial \ln \Delta w/\partial(1/T)$, where Δw is the weight of liquid imbibed, but in order to use them to evaluate the heat of dilution from equation (11), it is also necessary to estimate $(\partial \Delta G_0/\partial Q)_T$. A statistical expression for this quantity obtained from equation (4) may be used for the good swelling agents. In general, it is necessary to introduce the factor γ

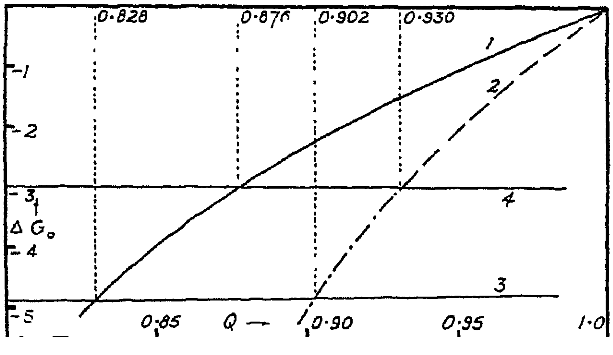


FIG. 6.—Theoretical and experimental dependence of Q_m on vulcanisation and extension.
1. ΔG_0 raw rubber (exptl.). 2. Do. (theoretical).
3. $-\partial \Delta G_0$ vulcanised rubber at $f = 0$. 4. Do. at $l/l_0 = 2$.

determined in the last section, as the ratio between the true and calculated values of $(\partial \Delta G_0/\partial Q)_T$. From equation (11), introducing the factor γ, and dividing by v_r^2 gives

$$\frac{\Delta H_0}{v_r^2} = \frac{\gamma R}{Q_m + 1}\left(-\frac{\partial \ln \Delta w}{\partial(1/T)}\right)\left\{1 + (1 - 2\mu)Q_m - \frac{\rho_r V_0}{3M_c}Q_m(Q_m + 1)^{5/3}\right\}, \quad (12)$$

where Q_m is the mean maximum swelling over the range of temperature considered and μ is determined by equation (6). The application of this equation to a single rubber in a range of liquids is illustrated in Table IV, values of γ being read from Fig. 5. Very similar results have been obtained for other rubbers.

TABLE IV.—HEAT AND ENTROPY OF DILUTION FROM TEMPERATURE DEPENDENCE OF SWELLING. (RUBBER B.)

Liquid.	Q_m.	$-\frac{\partial \ln \Delta w}{\partial(1/T)}$.	$\frac{\Delta H_0}{v_r^2}$. Cal./mole Liquid.	
			Eqn. (12).	Calorimetric[12] ($Q = 0$).
Carbon tetrachloride	6·90	− 125	− 105	− 255
Chloroform	6·81	− 185	− 130	− 450
Carbon disulphide	5·87	35	15	95
Toluene	5·23	30	20	15
Cyclohexane	5·05	75	50	
Benzene	4·96	235	125*	310*
n-Butyl *n*-butyrate	3·25	100	65	
n-Heptane	2·90	75	35	135
di-*iso*propyl ketone	2·82	330	160	
n-Pentane	2·58	− 115	− 45	
n-Butyl acetate	2·39	75	120	
n-Propyl acetate	1·63	620	205	
Ethyl acetate	0·84	1010	230	
Methyl ethyl ketone	0·62	1000	215	
Methyl acetate	0·30	905	155	
Acetone	0·18	925	120	950

* Direct measurement on raw rubber[4] gives $\Delta H_0/v_r^2 = 310$ at $Q = 0$; 168 at $Q = 4·96$.

Several features of this table call for comment. In the first place the results for the last two members of the series are very largely determined by γ, and it is clear from Fig. 5 that this may be quite seriously in error. For the next three members there are direct experimental measurements of γ; for the remainder $\gamma = 1$ is likely to be very nearly accurate.

At first sight the values of $\Delta H_0/v_r^2$ estimated here, do not agree at all well with those based on calorimetric data (compare columns 4 and 5). The discrepancy is always in the direction making the numerical value calculated from the temperature coefficient the smaller of the two. Now it has been shown by direct measurement[4] in the case of raw rubber and benzene that $\Delta H_0/v_r^2$ falls as the benzene content increases, and the value of *ca.* 168 cal./mole found at $Q = 4·96$ is in fair agreement with the estimate of 125 obtained here.

For this system the calorimetric estimate, which was based on the heats of mixing of benzene with dihydromyrcene,[12] is equal to the experimental value[4] at $Q = 0$. Since both columns 4 and 5 are essentially experimental their disagreement is evidence that $\Delta H_0/v_r^2$ is not independent of the degree of swelling, i.e. that even if μ is independent of composition, the " constants " A and B of equation (2) are not.

Conclusions.

The general conclusion which emerges from these experiments is that the simple equation (1) proposed by Huggins is more limited in application than has been previously supposed. It gives correctly the isothermal dependence on composition of the free energy of a polymer-liquid mixture whose heat of mixing is small, μ being treated as an empirical parameter. It is therefore possible to calculate the colligative properties of the mixture at a fixed temperature with fair precision using only the one adjustable parameter. Since the independence of μ on v_r is found to arise only from an approximate cancellation of the concentration dependence of A and B (equation 2), the equation cannot be safely used to discuss heats and entropies of dilution. If there is an appreciable heat of mixing, equation (1) is no longer strictly accurate even at a fixed temperature. As the polymer approaches saturation, the dependence of ΔG_0 on composition becomes considerably less than expected. The direct consequence of this is that theoretical treatments based on equation (1) underestimate the extent to which the swelling capacity of a polymer depends on external conditions or structural changes.

Discussion of the origin of these anomalies requires a critical examination of the basis of equation (1) and of other statistical theories of polymer solutions. This lies beyond the scope of the present paper, and will be considered in a further communication.

Summary.

Experimental data are reported on the effects of vulcanisation, extension and temperature on the swelling of rubber. These, although in qualitative agreement with a simple equation derived from a statistical theory of polymer solutions, reveal quantitative discrepancies which are important in discussions of two-phase equilibria.

The author's thanks are due to Miss P. Garnham for experimental assistance, to Mr. S. C. Stokes for the preparation of the rubber samples, and to Mr. T. A. Sharpley for the " modulus " determinations. The work described in this paper forms part of the programme of fundamental research on rubber undertaken by the Board of the British Rubber Producers' Research Association.

British Rubber Producers' Research Association,
48 Tewin Road,
Welwyn Garden City, Herts.

APPENDIX I.

Thermodynamic Relations.

(i) Dependence of Q_m on Extension.—Considering a stretched, swollen rubber maintained at constant temperature and pressure; by mathematical identity

$$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_l = -\left(\frac{\partial \Delta G_0}{\partial l}\right)_Q \left(\frac{\partial l}{\partial Q}\right)_{\Delta G_0} \quad . \quad . \quad . \quad (13)$$

and

$$\left(\frac{\partial \Delta G_0}{\partial l}\right)_Q = \left(\frac{\partial}{\partial l}\right)_Q \left(\frac{\partial G}{\partial n_0}\right)_l = \left(\frac{\partial}{\partial n_0}\right)_l \left(\frac{\partial G}{\partial l}\right)_Q = \left(\frac{\partial f}{\partial n_0}\right)_l, \quad . \quad (14)$$

where f is the tension in the rubber and n_0 the number of moles of liquid in the rubber.

Further,

$$n_0 = A_0 l_0 Q / V_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

whence,

$$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_l = -\frac{V_0}{A_0 l_0}\left(\frac{\partial f}{\partial Q}\right)_l\left(\frac{\partial l}{\partial Q}\right)_{\Delta G_0}. \quad . \quad . \quad . \quad (16)$$

A particular case of equation (16) is obtained by putting $\Delta G_0 = 0$, when $Q = Q_m$ giving

$$l_0\frac{\partial Q_m}{\partial l} = -\frac{V_0}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l \Big/ \left(\frac{\partial \Delta G_0}{\partial Q}\right)_l \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where the terms on the right are to be given their values at $Q = Q_m$.

(ii) **Swelling at Constant Tension and at Constant Length.**—Equation (9) may be used to derive the limiting value of $(\partial \Delta G_0/\partial Q)_l$ at $f = 0$. To obtain from this the slightly different function $(\partial \Delta G_0/\partial Q)_{f=0}$, we use the identity

$$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_{f=0} = \left(\frac{\partial \Delta G_0}{\partial Q}\right)_l + \left(\frac{\partial \Delta G_0}{\partial l}\right)_Q\left(\frac{\partial l}{\partial Q}\right)_{f=0}. \quad . \quad . \quad . \quad (17)$$

Now at $f = 0$, $l = l_0(Q+1)^{\frac{1}{3}}$, so that

$$\left(\frac{\partial l}{\partial Q}\right)_{f=0} = \tfrac{1}{3} l_0 (Q+1)^{-\frac{2}{3}}. \quad . \quad . \quad . \quad . \quad . \quad (18)$$

Combining this with the value of $(\partial \Delta G_0/\partial l_0)_Q$ obtained from (14) and (15), and of $(\partial \Delta G_0/\partial Q)_l$ from (9), the final result is

$$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_{f=0} = -\frac{V_0}{A_0}\left(\frac{\partial f}{\partial Q}\right)_l\left\{\frac{1}{l_0(\partial Q_m/\partial l)} - \tfrac{1}{3}(Q_m + 1)^{-\frac{2}{3}}\right\} \quad . \quad (10)$$

where the quantities on the right are to be given their values at $f = 0$, $Q = Q_m$.

(iii) **Temperature Dependence of Q_m.**—Considering a swollen rubber at constant pressure, and free from mechanical restraint; by identity

$$\left(\frac{\partial \Delta G_0}{\partial Q}\right)_T = -\left(\frac{\partial \Delta G_0}{\partial T}\right)_Q \Big/ \left(\frac{\partial Q}{\partial T}\right)_{\Delta G_0}. \quad . \quad . \quad . \quad (19)$$

If w_0 is the weight fraction of liquid in the rubber, the entropy of dilution ΔS_0 is given by the well-known relation

$$\begin{aligned} -\Delta S_0 &= (\partial \Delta G_0/\partial T) w_0 \\ &= \left(\frac{\partial \Delta G_0}{\partial T}\right)_Q + \left(\frac{\partial \Delta G_0}{\partial Q}\right)_T\left(\frac{\partial Q}{\partial T}\right)_{w_0} \quad . \quad . \quad . \quad (20) \end{aligned}$$

by identity. If α_0 and α_r are the coefficients of thermal expansion of liquid and rubber, it is easily seen that

$$\left(\frac{\partial Q}{\partial T}\right)_{w_0} = (\alpha_0 - \alpha_r)Q. \quad . \quad . \quad . \quad . \quad . \quad (21)$$

Substituting in (26) and combining with (25)

$$\Delta S_0 = \left(\frac{\partial \Delta G_0}{\partial Q}\right)_T\left\{\left(\frac{\partial Q}{\partial T}\right)_{\Delta G_0} - Q(\alpha_0 - \alpha_r)\right\}. \quad . \quad . \quad (22)$$

In the particular case of maximum swelling, $\Delta G_0 = 0$, $Q = Q_m$ and

$$\Delta H_0 = T\Delta S_0 = T\left(\frac{\partial \Delta G_0}{\partial Q}\right)_T\left\{\frac{\partial Q_m}{\partial T} - Q_m(\alpha_0 - \alpha_r)\right\}. \quad . \quad . \quad (23)$$

In using experimental data obtained by weighing, it is simpler to evaluate the temperature coefficient of the weight increase Δw, and replace equation (23) by the equivalent expression

$$\Delta H_0 = T\Delta S_0 = -\frac{Q_m}{T}\left(\frac{\partial \Delta G_0}{\partial Q}\right)_{T,\, Q=Q_m} \cdot \frac{\partial \ln \Delta w}{\partial (1/T)}. \quad . \quad . \quad (11)$$

APPENDIX II.

COMPOSITION OF VULCANISED RUBBERS.

	A, A′.	B, B′, B″.	C.	D.
Smoked sheet . . .	100	100	100	100
Sulphur	2	3	3	5
Selenium (Vandex) . .	—	—	2	—
Zinc oxide . . .	2	2	2	2
Stearic acid . . .	0·5	0·5	0·5	0·5
Nonox	0·5	0·5	0·5	0·5
M B T S	1	—	—	—
Z D C	—	1	1	1

APPENDIX III.

" CURE " OF RUBBERS.

	A.	A′.	B.	B′.	B″.	C.	D.
Cure (Min.) . .	30	30	60	15	30	60	60
at (°c.) . .	142	142	100	142	142	100	100
Modulus (300 %) (kg./cm.^{2}) .	9	—	24	14	13	28	30
10^{-3} M_C . .	8·0	11·1	6·4	7·9	8·3	$5{\cdot}1_5$	4·5
10^{-4} Modulus × M_C	(72)	—	155	110	110	145	135

A and A′ differed in the amount of milling received during mixing; A had a bare minimum and was of Mooney viscosity 75; A′ was heavily milled to a Mooney viscosity of 15.

GENERAL DISCUSSION *

Dr. R. F. Tuckett (*Welwyn*) said: Gee's results on the swelling of stretched rubber show that the Flory-Huggins equation with its single arbitrary constant μ is not adequate in the phase separation region where ΔH_0 is relatively large. This is not unexpected, and it is of interest to examine a system whose apparent properties correspond very closely to the simple model on which the whole theory is based. This is so if the polymer unit and the solvent molecules are geometrically and chemically almost identical so that the heat of dilution is zero; in such a system $B = 0$ and μ becomes independent of temperature, being a function of z only, where z is the co-ordination number of the lattice on which the molecules are arranged. As $\Delta G_0 = RT \ln p/p_0$, then, for long chains, the relation between p/p_0 and w_r, the weight fraction of polymer, can be calculated from Guggenheim's equation as a function of z only.[1] For such an " ideal " polymer-solvent system, the following theoretical figures for p/p_0 are obtained for three values of w_r.

z.	$w_r = 0{\cdot}8$.	$w_r = 0{\cdot}65$.	$w_r = 0{\cdot}50$.
	$p/p_0 =$	$p/p_0 =$	$p/p_0 =$
6	0·51	0·73	0·87
8	0·49	0·71	0·85
∞	0·45	0·67	0·82

$z = \infty$ corresponds to the simple Flory-Huggins expression (Gee's equation (1)) and $z = 6$ seems to be the smallest value of z which is physically satisfying. The table shows, therefore, the plausible limits inside which p/p_0 should fall.

---

* On preceding paper.

[1] Guggenheim, *Proc. Roy. Soc. A*, 1944, 183, 203.

The experimental polymer-solvent systems which conform most closely to the theoretical requirements are those in which the solvent is either the monomer or the saturated monomer of the second polymeric component. The system polyvinyl acetate-ethyl acetate is especially suitable as its heat of dilution has been shown to be zero by direct measurement.[2] Walling, Briggs and Mayo [3] also give kinetic and activity results from which the p/p_0-w_r relation can be obtained for the polystyrene-styrene system. Results, which will be reported more fully elsewhere, are as follows :

$\frac{w_r}{p/p_0}$.	·80.	0·65.	0·50.
Polyvinyl acetate-ethyl acetate . . .	0·59	0·83	0·93
Polystyrene-styrene .	0·60	0·80	0·93

These values fall quite outside those of the previous table; to obtain agreement, z values of the order of three are necessary, which seem physically impossible. Similar discrepancies are found in the osmotic pressure region and have been explained by Flory [4] but his modified theory is restricted to very dilute solutions. It seems unlikely that any end effects are significant for really long chains, and the results suggest that the whole theory requires re-examination.

Prof. E. A. Guggenheim (*Reading*) said : Just as the experimenter, in choosing the materials and details of his research, is largely guided by matters of convenience, so the theoretician in his research wisely begins by considering the simplest possible models before attempting to tackle more complicated ones. It is important to bear this in mind when making a comparison between experimental data and theoretical formulæ.

Up to the present, all the theoretical formulæ are based on the following assumptions (1) Each type of molecule occupies a definite integral number of sites on a lattice. (2) The energy of interaction is the sum of terms contributed by each pair of neighbouring sites. (3) The energy of interaction contributed by each pair of sites depends only on what species occupies each of the sites and not on which element of a molecule occupies either site.

Assumptions (1) and (2) are probably good enough for the purpose of comparison with the available experimental data; but in my opinion assumption (3), which means regarding each molecule as completely homogeneous, is seriously wrong. Unpublished experimental data from Professor Bronsted's laboratory in Copenhagen show that two different normal paraffins have small, but far from negligible, heats of mixing. Since the middle portion of two different normal paraffins are indistinguishable, it seems obvious that the heat of mixing must be due to the end groups. It is therefore my opinion that the next stage in improving the theory should be a study of arrangements on a lattice of molecules with distinguishable elements and I hope to attack this problem in the near future.

From these considerations it would at first appear that the models already investigated would at least be sufficient in the case of zero energy of mixing. This, however, is not necessarily true for it is conceivable that the energy of interaction between two molecules may be non-zero, but may change sign when one of the molecules is reversed. In this case, a resultant zero energy of mixing may well be due to a balance between equal positive and negative contributions from various pairs of sites.

Dr. M. Magat (*Paris*) said : I should like to comment on the determination of the heat of mixing from swelling experiments and to give a possible reason for the disagreement between the $\Delta H/v_r^2$ values so obtained with the calorimetric values for $Q = 0$. In replacing the differential $\partial \ln \Delta w/\partial(1/T)$ by the difference, one assumes $\Delta H_{\text{mix.}}$ to be independent of temperature, which is definitely not the case. Doty and Zable,[5] for instance, have found for polyvinylchloride that in a few cases $\Delta H_{\text{mix.}}$ increases with temperature which can be explained by assuming either that the co-ordination number or that the heat of mixing increases with T. Only the second alternative seems to be a reasonable one. Similarly, Desmaroux[6] has found that in certain cases nitrocellulose is more soluble at lower temperatures than at higher ones, which also points to a substantial increase of ΔH_{mix}. This strong variation of $\Delta H_{\text{mix.}}$ and its increase with temperature in certain cases, can be predicted from the Scatchard-Hildebrand formula.

---

[2] P. Meares (private commnnicatoin).

[3] Walling, Briggs and Mayo, *J. Amer. Chem. Soc.*, 1946, 68, 1145.

[4] Flory, *J. Chem. Physics*, 1945, 14, 453.

Dr. P. M. Doty (*Cambridge*) said : Dr. Gee has quite adequately demonstrated by several methods the shortcomings of the Flory-Huggins equations in the region of two-phase equilibria. It may be of interest, however, to point out a recent experiment which affects somewhat his results shown in Fig. 1 and 2.

Osmotic pressure measurements on solutions of polystyrene containing small amounts of divinyl benzene show that μ increases with the amount of cross-linking, at least for fairly indifferent solvents. Boyer and Spencer have demonstrated that this effect is also present in three-dimensional gels of polystyrene-divinyl benzene. If this is taken into account the lines of constant μ in Fig. 1 and 2 would be modified in the direction of closer agreement.

A failure of the Flory-Huggins equation in the region of two-phase equilibrium was noted in our study of the polyvinyl chloride-dioxane system. Here values of μ as high as 0·515 were observed for solutions, whereas the critical value beyond which phase separation should occur is calculated to be about 0·504.

Nevertheless, we should really not speak of these observations as failures of the Flory-Huggins theory for by its nature it cannot be expected faithfully to represent in a quantitative manner the detailed behaviour of polymer solutions. The recent studies by Flory and by Zimm have shown the way to overcome its greatest defect—the assumption of uniform concentration of polymer on a molecular scale—but at the same time emphasise how difficult it will be to predict accurately the relative contribution of heats and entropies to μ.

Dr. J. J. Hermans (*Utrecht*) said : Dr. Gee finds that the change in degree of swelling on extension conforms much better to theory if the solvent is a good solvent and the degree of swelling accordingly is high, than if swelling is low. From Dr. Gee's paper and the subsequent discussions it follows that this discrepancy between theory and experiment is due primarily to the shortcomings of the thermodynamic theory. Nevertheless, it is important to realise that there is an additional effect which I believe is not accounted for in the formulæ used by Gee.

If a piece of rubber is vulcanised in the dry state, we may assume that the statistical coiling in that state conforms to the well-known requirements of Kuhn's theory. Let us call this the state of " normal " coiling. If this rubber swells in a liquid, the molecular chains are stretched, and coiling must become *sub*normal. In the reverse case, where the volume of the gel is less than that where coiling is normal, there will be *super*coiling. This effect introduces a " volume factor " into the equation for the retractive force of the network. In the case of supercoiling, for instance, we may say that for a given macroscopic elongation of the network the effective elongation of the chain molecule is *less* than if coiling were normal. Supercoiling would therefore displace the theoretical curves of Fig. 3 and Fig. 4 in Gee's paper toward lower extension, i.e. higher Q_m, and this displacement would be larger in Fig. 4, where the degree of swelling is low, than in Fig. 3 where swelling is high.

---

[5] Doty and Zable, *J. Polymer Science*, 1946, 1, 90.

[6] Desmaroux, *Mémorial des Poudres*, 1917, p. 168.

Dr. J. R. Scott (*Croydon*) (*communicated*) : It has long been known[7] that there is an approximate proportionality between the swelling of vulcanised natural rubber in benzene and its " hardness number," expressed as the depth of indentation by a steel ball under a constant load, except that this relation breaks down with rubbers hardened by incorporation of large proportions of filler. Using Gee's relation (6) between swelling and M_c, the relation between M_c and elastic modulus for small extensions,[8] and the known relation between this modulus and the hardness number[9] applying corrections for the effect of finite specimen thickness on the hardness measurement, the following calculated hardness numbers corresponding to various degrees of swelling are obtained, and are compared below with values observed by experiment.

Swelling in benzene (Q_m)	2	3	4	5	6
Hardness number (mm./100), calc. . . .	63	99	135	172	210
,, ,, ,, obs. . . .	45	95	137	167	195

(hardness measured with ¼-in. diam. ball and 1 kg. load on 1 cm. thick specimen ; " observed " values are approximate because rubbers of the same swelling vary in hardness number by ± 10).

The calculated hardness number shows fairly close proportionality to the swelling, and except for the lowest Q_m value, the agreement between calculated and observed values is very good in view of the simplifying assumptions involved in the relations used.

Dr. G. Gee (*Welwyn Garden City*) (*in reply*) said : Instead of trying to reply to each query separately, I think it will be best to try and summarise briefly my present views on the interpretation of thermodynamic data on polymer solutions.

In discussing the significance of the various ways in which Huggins' equation (1) fails to agree with experiment, it is first necessary to examine the basis of the equation. The first point to be noted is that it results from combining expressions separately deduced for the entropy ΔS_0 and heat ΔH_0 of dilution, it being assumed that mixing is random. The effect of a finite heat of mixing on ΔS_0 has been explicitly calculated by Orr[10] and Guggenheim,[11] employing the same lattice model of the system as Huggins. Numerical evaluation of their somewhat complex equations leads to the conclusion that ΔS_0 should not be very dependent on ΔH_0, so that no serious error should arise from assuming mixing to be random.

The second point at which Huggins' treatment may be criticised is that the heat of dilution assumed is inconsistent with the model used in calculating ΔS_0. The equation $\Delta H_0 = \alpha v_r^2$ is known to hold for many liquid mixtures and is the form to be expected for a random mixture of molecules of equal size. The lattice model of a polymer system requires $\Delta H_0/v_r^2$ to diminish approximately in the ratio of 1 to $\left(1 - \frac{2}{Z}\right)^2$ as v_r goes from 1 to 0, Z being the co-ordination number of the lattice, and most of the decrease occurring between $v_r = 0{\cdot}5$ and $1{\cdot}0$.

Experimentally[12] $\Delta H_0/v_r^2$ is found to be nearly constant for the systems, squalene $(C_{30}H_{50})$ — benzene or — heptane, in disagreement with the deduction from a lattice model. The dependence of $\Delta H_0/v_r^2$ on v_r found for rubber and benzene[13] is not of the expected form and is believed to have another origin (see below). From the above considerations it appears that the lattice model gives the wrong form for the heat of dilution, and hence also for ΔG_0. The success of Huggins' equation in giving the correct form of ΔG_0 for a number of systems must certainly arise from a compensation of errors.

---

[7] Scott, *Trans. Inst. Rubber Ind.*, 1929, **5**, 95.

[8] Wall, *J. Chem. Physics*, 1942, **10**, 485 ; Flory and Rehner, *ibid.*, 1943, **11**, 512.

[9] Scott, *Trans. Inst. Rubber Ind.*, 1935, **11**, 224.

[10] Orr, *Trans. Faraday Soc.*, 1944, **40**, 320.

[11] Guggenheim, *Proc. Roy. Soc. A.*, 1944, **183**, 213.

[12] Ferry, Gee and Treloar, *Trans. Faraday Soc.*, 1945, **41**, 340.

[13] Gee and Orr, *ibid.*, 1946, **42**, 507.

A way of resolving this dilemma is afforded by Guggenheim's extension[14] of the statistical treatment to a modified lattice model in which each molecule of liquid occupies several sites. Miller[15] has recently shown that these need not be in the form of a chain. A detailed analysis of the effects of this change in the model will be published elsewhere, but broadly speaking it is found to be equivalent to increasing the value assigned to the lattice co-ordination number in the previous model. Indeed to a fair approximation the results are now the same as those obtained by assuming $Z = \infty$. Thus $\Delta H_0/v_r^2$ becomes more nearly constant and the entropy of dilution reduces very nearly to the simple form originally deduced by Flory :[16]

$$\Delta S_0 = -R(\ln v_0 + v_r).$$

Guggenheim's treatment therefore provides the first satisfactory basis for Huggins' equation, and departures from the equation may properly be ascribed to defects in the model adopted by Guggenheim. One defect which is now generally recognised is the failure to allow for the kinking of the individual molecule, which results in an increased number of polymer-polymer contacts at low values of v_r. It has been suggested elsewhere[13] that this is responsible for the whole of the observed concentration dependence of $\Delta H_0/v_r^2$ for the system rubber and benzene. A compensating reduction in ΔS_0 leaves ΔG_0 in close agreement with Huggins' equation.

Another type of departure which may occur is associated with volume changes on mixing; these may arise from directed intermolecular forces, or from purely geometrical factors. A striking example of this latter effect is furnished by rubber-*n*-pentane. The negative value of ΔH_0 for this system (Table IV of my paper) has been shown[17] to result from a large contraction on mixing. In general, the effect of volume changes will be mainly confined to changing the relative values of the temperature dependent and independent parts of μ, leaving μ comparatively unchanged.

Neither of these phenomena account for the systematic failure of Huggins' equation when applied to poor swelling agents at concentrations close to that of the two-phase equilibrium. The only explanation which seems consistent with the facts is that as saturation is approached, a state of non-random mixing arises, in which clusters of the small molecules are formed. This may be seen very clearly by comparing the two values of $\Delta H_0/v_r^2$ for rubber-acetone given in Table IV. These figures suggest strongly that the probable number of rubber-acetone contacts made by an acetone molecule entering the system at the saturation point is some eight times smaller than would be expected from random mixing. The present model fails completely to predict that such a state of non-random mixing would be stable. Indirect evidence of non-random mixing in liquids has been drawn by Schallamach[18] from measurements of the dielectric loss of mixtures of dipolar liquids in which the heat of mixing would be expected to be large. The problem is therefore primarily one of the liquid state, and is not essentially connected with polymer solutions.

Dr. G. Salomon (*Delft*) (*communicated*): The fact that the change of a swelling equilibrium under influence of mechanical deformation and temperature can be expressed in terms of statistical thermodynamics indicates by itself a considerable advance in theory. Agreement with experiments is limited by the necessary assumption that the strength of molecular interactions, which are involved in the equilibrium, *i.e.* the contact between liquid-liquid, liquid-polymer and polymer-polymer is independent of such parameters as cross-linking and extension. Fig. 5 and Table IV of Gee's paper outline the narrow limits between which such an assumption may be valid for natural rubber.

---

[14] Guggenheim, *Proc. Roy. Soc. A.*, 1944, **183**, 203.

[15] Miller, *Nature*, 1946, **157**, 842.

[16] Flory, *J. Chem. Physics*, 1942, **10**, 51.

[17] Gee (unpublished work).

[18] Schallamach, *Trans. Faraday Soc.* (1946, **42A**).

We made an extensive study of swelling equilibria using a series of elastomers of increasing polarity and a large variety of solvents.[19] Our experiments indicate that molecular interaction becomes the more complex the more polar groups are introduced in either the polymer or the liquid. Generally speaking the application of equations (1), (7) and (12) will be limited in the same way as the calculation of swelling equilibria from cohesional energy densities, which becomes less satisfactory with increasing complexity of intermolecular forces either in the polymer or in the liquid.

Such limitations in the possible application are *not* due to an imperfection in Gee's theory, but to the perplexing and little understood correlation between molecular structure and intermolecular forces. Not only varies this interaction between the second and the sixth power of intermolecular distance, but it is also depending on the degree of free rotation as well as on the shape of the molecular surface. Moreover the distribution of charges is also a function of the molecular surface as outlined by van Arkel and Staverman. As a consequence, molecular interaction in a system, *e.g.* polystyrene-benzene, is fundamentally different from that in the molecular analogue, ethylbenzene-benzene.

The study of liquid mixtures affords numerous tedious measurements, while swelling equilibria and precipitation points are easily measured and provide the same type of information. The tremendous sensitivity of polymer solutions and gels against small changes in molecular structure allows as well a study of the properties of liquids,[20] using a polymer as a base of reference as the characterisation of the polymer by its interaction with liquids. The quantity γ in Fig. 5 and variations in the constants A and B of equation (2) may in future prove useful for such a purpose.

We have studied the influence of extension on the swelling of fully vulcanised samples (i.e. more cross-linked than those used by Gee) and have reached somewhat different results. The increase in swelling was distinctly smaller than the one illustrated by Fig. 3 of Gee's paper and moreover it proved to be impossible to keep the rubber under more than 150 % extension even in a weak swelling agent such as alcohol and acetone, an effect quite comparable to the puzzling behaviour of polythene in non-solvents.[21] It is quite possible that the samples used by Gee exhibit a certain amount of plastic flow under the conditions of Fig. 4. The conclusion that the tension of swollen rubber is independent of the nature of the liquid is corroborated by other measurements from this laboratory.[22]

Prof. E. C. Baughan (*Shrivenham*) said: I would like to make some comments on Dr. Gee's paper based on some work done by Prof. Brønsted and myself in 1937—publication has been delayed by the war. In this work thin films of polystyrene were weighed electromagnetically in organic vapours at measured pressures; the experiments were made at 20° C. and by using films about 0·02 mm. thick hysteresis could be reduced to minor importance. The observations covered volume fractions of solvent from about 0·1 to 0·5, and ranged from vapours of true solvents in which polystyrene will *dissolve* to those of liquids in which (at 20°) it will only *swell*. To apply the Huggins equation to such data it is necessary to *assume* no volume change on mixing. But when this is done it is found that the equation describes the facts very well with the following single exception; that for practically saturated *swelling* vapours, the amount of vapour picked up by the film was greater than the predicted value from a Huggins equation fitting lower pressures. This deviation was *not* observed for *dissolving* vapours. This observation confirms those made by Gee.

---

[19] Salomon and van Amerongen, *J. Polymer Sci.* (in press).

[20] Brønsted, *C.R. Trav. lab. Carlsberg*, 1938, **22**, 99.

[21] Richards, *Trans. Faraday Soc.*, 1946, **47**, 27.

[22] Wildschut, *Technological and Physical Investigations on Natural and Synthetic rubbers* (Elsevier, Amsterdam, 1946).

The μ values appropriate are given in the following list. *Dissolving vapours*: benzene 0·43; toluene 0·47; *m*-xylene 0·41; CCl_4 0·38; dioxane 0·56; diethylketone 0·67; *n*-butyl acetate 0·82. *Swelling vapours*: *n*-dipropyl ether 0·89; cyclohexane 0·89; nitromethane about 1·75. For benzene, toluene, CCl_4, and dioxane, μ — values can also be obtained for very dilute solutions from the osmotic data by other authors,[23] the values obtained being, benzene 0·45, toluene 0·44, CCl_4 0·46, dioxane 0·47.

The values for benzene and toluene are practically the same in the two sets of results; those for CCl_4 and dioxane are significantly different as must also be those for diethyl ketone and *n*-butyl acetate since $\mu > \frac{1}{2}$ implies that the liquid can be in equilibrium with a gel, i.e. that the liquid only *swells* while, in fact, these liquids *dissolve* polystyrene.

This work supports therefore Gee's contentions in that

(*a*) a complete isotherm characterised by one μ-value in Huggins' equation appears the exception rather than the rule—though it appears in some systems where the high, and low, molecular components are chemically very similar.

(*b*) The equation fails for "swelling" vapours in the neighbourhood of their saturation vapour-pressure.

It is hoped to publish work soon in more detail.

---

[23] Schulz, *Z. physikal. Chem. A*, 1936, **176**, 317, Dobry and Schwob, *Bull. soc. chim. France*, *T*, 1936, **2**, 1790.

THE EFFECT OF SOLVENT-TYPE AND TEMPERATURE ON THE SPECIFIC VISCOSITIES OF POLYSTYRENE SOLUTIONS.

By Turner Alfrey, J. David Justice and Samuel J. Nelson.

Received 24*th July*, 1946.

Many flexible linear polymers have been shown to exhibit higher intrinsic viscosities in good solvents than in poor solvents. This fact is in harmony with the currently accepted theories of polymer solutions; in a good solvent, the molecule should be significantly "swollen"—i.e. extended—while in a poor solvent "deswelling" or curling up should take place.

Gee[1] and Ewart[2] have pointed out that a close correlation exists between the swelling of a cross-linked polymer and the intrinsic viscosity of a (chemically similar) linear polymer in a given solvent. Furthermore, in both cases the degree of swelling is dependent upon the partial molal heat of dilution, $\overline{\Delta H_1}$. The partial molal heat of dilution in turn can be related (in simple cases) with the cohesive energy densities of the solvent and solute, or with an interaction constant α which is the square root of the cohesive energy density.[3] By definition,

---

[1] Gee, *British Rubber Producers Research Assn.*, Publ. No. 10, The Physical Chemistry of Rubber Solutions; *India Rubber J.*, 1935, **108**, 349; *Trans. Faraday Soc.*, 1944, **40**, 463; *ibid.*, 1944, **40**, 480.

[2] Ewart, *Recent Advances in Colloid Science*, vol. 2 (Interscience Publishers Inc., New York, 1946), p. 197.

$$\text{Cohesive energy density} = \alpha_1^2 = \frac{L_1 - RT}{V_1}$$

where L_1 is the molar heat of vaporisation and V_1 the molar volume of the solvent. The interaction constant of the polymer, α_2, can be estimated from swelling data. In ideal cases the partial molal heat of dilution is given by the equation

$$\overline{\Delta H_1} = K \,.\, V_1(\alpha_2 - \alpha_1)^2 V_2^2$$

where K is a small constant approximately equal to unity, V_1 the molar volume of the solvent, and V_2 the volume fraction of the polymer. This equation states that the more nearly the interaction constant of the solute is equal to that of the solvent, the lower will be the heat of mixing,[3] and indicates that for each polymer there is a solvent which will be " best," and in which the polymer will have the highest intrinsic viscosity. This value for intrinsic viscosity should remain nearly constant with increasing temperature. If the interaction constant of the solvent is significantly different from that of the polymer, the intrinsic viscosity will be lower, and will increase with increasing temperature. This should also mean that if a solvent is " poor " because it is too polar, its properties should be improved by the addition of a non-polar material; if too non-polar, it should be improved by the addition of a polar material.

Gee[1] has considered these factors for natural and synthetic rubbers in terms of cohesive energy densities, and has shown that it is possible to obtain solvents for Neoprene GN, Buna S, and Buna N from mixtures of two non-solvents in cases where the interaction constants of the non-solvents lie on opposite sides of the interaction constant for the polymer. Brønsted and Volquartz[4] investigated several polymers in mixed solvents, and found, for example, that polystyrene can be dissolved in a mixture of acetone and *n*-propyl laurate.

The effect of mixtures of solvents on intrinsic viscosity has not been as extensively studied as the effect of pure solvents. Alfrey, Bartovics and Mark[5] have measured η_{sp}/c at a given concentration in mixtures of a solvent and non-solvent for rubber and polystyrene. The specific viscosity decreased to a limiting value at the precipitation point which was approximately the same in all cases. This fact was interpreted as indicating that a solvent mixture just poor enough to cause an appreciable inter-molecular agglomeration of polystyrene molecules (precipitation) would also cause a certain definite amount of intra-molecular agglomeration (curling up of individual molecules).

Ewart[2] presents a generalised relationship between the interaction constant of the solvent, and the intrinsic viscosity in which he indicates that the $[\eta]$ will decrease as the absolute value of $(\alpha_2 - \alpha_1)$ increases. In the present investigation the general type of behaviour described above was observed. The correlation of specific viscosity with cohesive energy density was found, however, to be only semi-quantitative, with particularly large deviations in the case of polar, asymmetric solvents.

Experimental Procedure.

(A) Preparation of Polymer.—The polystyrene used in these investigations was prepared by Dr. W. P. Hohenstein. Styrene monomer (Dow N-99) was washed with a 10 % solution of sodium hydroxide, and with water. The monomer then stood overnight over anhydrous calcium chloride, and was distilled under reduced pressure. The inhibitor-free styrene was then permitted to polymerise, at room temperature, over a period of six months. The conversion was 27 %, as determined by precipitating the polystyrene with methanol, from a benzene solution of the reaction mixture.

---

[3] Scatchard, *Chem. Rev.*, 1931, **6**, 321.

[4] Brønsted and Volquartz, *Trans. Faraday Soc.*, 1940, **36**, 619.

[5] Alfrey, Bartovics and Mark, *J. Amer. Chem. Soc.*, 1942, **64**, 1557.

(B) Fractionation.—The polystyrene was fractionated by successive precipitation using methyl ethyl ketone as solvent and methanol as non-solvent. Two successive fractionations resulted in 8 fractions designated by the symbols JN 1 through JN 8. Fractions 1 through 5 were refractionated by Dr. Bruno Zimm by a series of recombinations and refractionations using the same solvent—non-solvent system. The five fractions obtained by this procedure are designated by the symbols JNZ 1 through JNZ 5.

(C) Determination of Intrinsic Viscosity.—Intrinsic viscosities were determined by extrapolation of plots of η_{sp}/c against c (concentration was expressed in grams per 100 ml. of solution). The measurements were corrected for concentration changes resulting from thermal expansion, and kinetic energy corrections were made on all determinations for which such corrections would be appreciable.

(D) Determination of Precipitation Points.—Precipitation points were determined by adding precipitant to 5 ml. of solution (0·5 g./100 ml.) in a 25 ml. Erlenmeyer flask. A micro-burette was used to measure the volume of precipitant. The end-point was the point at which the solution became quite cloudy, and was very sharp in all cases. After determining the precipitation point at 20° c., the polymer was redissolved by placing the flask in the 40° c. bath, and the determination repeated. The same process was repeated at 60° c.

Results and Discussion.

(A) Intrinsic Viscosities in Mixed Solvents.—Figure 1 shows the variation of intrinsic viscosity of fraction JNZ-2 with temperature and

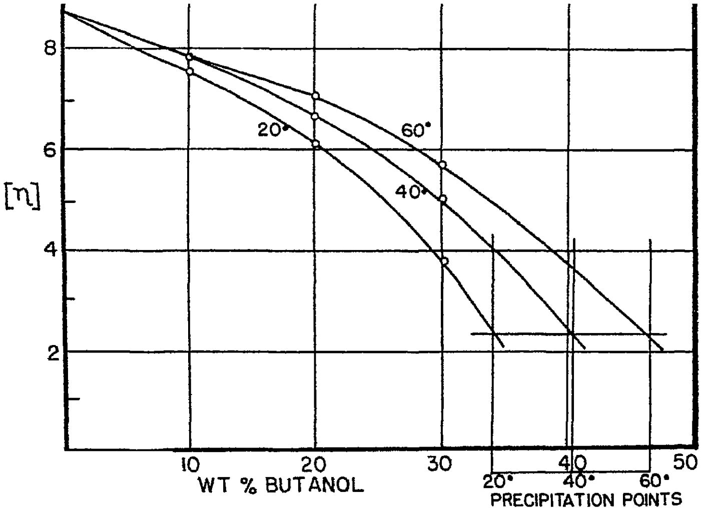


FIG. 1.—Intrinsic viscosity of fraction JNZ-2 in toluene-butanol.

with the composition of mixed solvent in the case of the system toluene-butanol. Increasing amounts of butanol in the solvent mixture results in a regular decrease in the intrinsic viscosity at 20°, 40° and 60° c. When corrected for concentration changes, the intrinsic viscosity in pure toluene is, within experimental error, independent of temperature. In the more unfavourable mixtures, however, the intrinsic viscosity increases markedly with an increase in temperature, as is to be expected. The curves extrapolate to an intrinsic viscosity of slightly over 2·0 at the precipitation points. In this case, the extrapolations are rather dubious. The data are consistent with the observation [5] that a definite intrinsic viscosity characterises the precipitation point, but are not sufficient to prove this observation. In Fig. 2 are shown similar curves for the system toluene-heptane. The results are qualitatively similar and it is significant that the curves again extrapolate to an intrinsic viscosity slightly greater than 2·0 at the precipitation point. In Fig. 3 are shown similar data for the system methyl ethyl ketone-heptane. Three significant features of these latter curves may be pointed out.

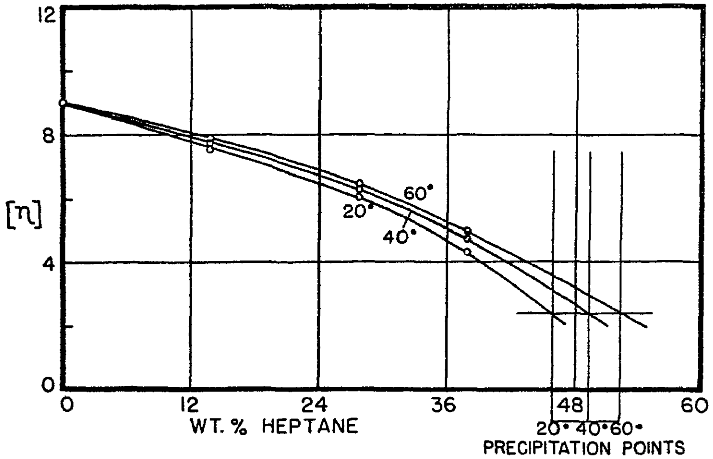


Fig. 2.—Intrinsic viscosity of fraction JNZ-2 in toluene-heptane.

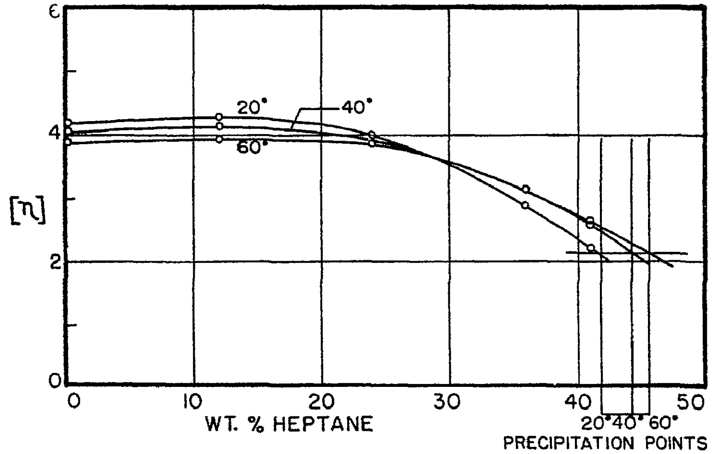


Fig. 3.—Intrinsic viscosity of fraction JNZ-2 in methyl ethyl ketone-heptane.

(1) Again the curves extrapolate to a specific viscosity of slightly greater than 2 at the precipitation points.

(2) Although pure methyl ethyl ketone is a poor solvent as indicated

by the low value of intrinsic viscosity, the intrinsic viscosity *decreases* with an increase in temperature. This rather anomalous behaviour is

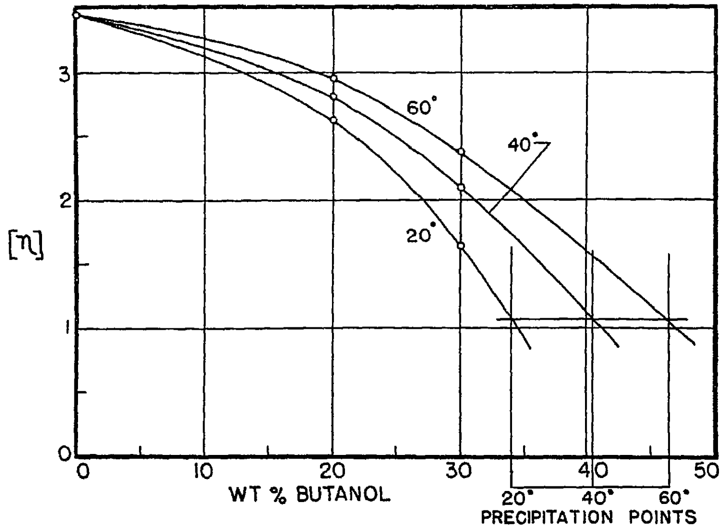


Fig. 4.—Intrinsic viscosity of fraction JN-7 in toluene-butanol. Precipitation points are for fraction JNZ-2.

presumably due to the fact that as the temperature rises, the calculated α of methyl ethyl ketone becomes increasingly different from the favourable α, as represented by toluene. Thus methyl ethyl ketone becomes

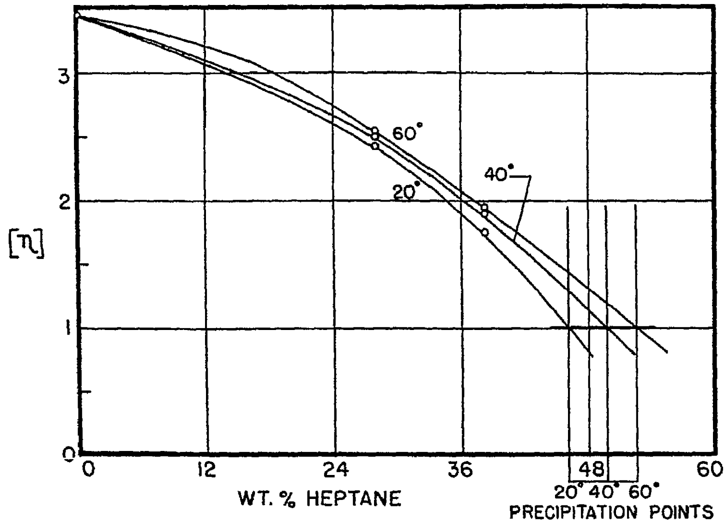


Fig. 5.—Intrinsic viscosity of fraction JN-7 in toluene-heptane. Precipitation points are for fraction, JNZ-2.

an increasingly unfavourable solvent as the temperature rises, and the intrinsic viscosity therefore decreases.

(3) In the two preceding cases, the intrinsic viscosity began to decrease with the first addition of non-solvent. The first 20 % of heptane, how-

ever, causes no decrease in the intrinsic viscosity of a polystyrene solution in methyl ethyl ketone. In fact, there is some indication of a small *increase* in intrinsic viscosity upon addition of a small quantity of heptane. This difference in behaviour is almost certainly connected with the fact that methyl ethyl ketone has a higher cohesive energy density than polystyrene while heptane has a lower cohesive energy density. Indeed, a naive application of the concept of cohesive energy density would actually predict a much *larger* deviation in this direction. One would expect some intermediate solvent mixture, possessing an average cohesive energy density close to that of polystyrene, to give an intrinsic viscosity equivalent to that observed in toluene. The fact that such a pronounced maximum does not occur in Fig. 3 is an indication of the (quantitative) failure of the cohesive energy density concept in the case of methyl ethyl ketone.

Fig. 4 and 5 show similar results for fraction JN-7.

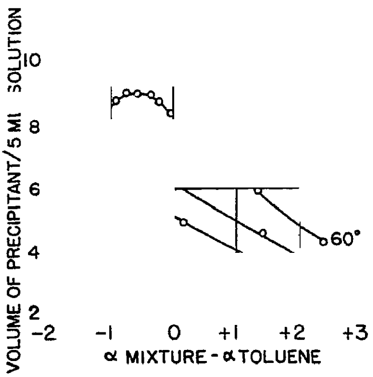


FIG. 6.—Volume of mixed precipitant (heptane-butanol) required to cause precipitation.

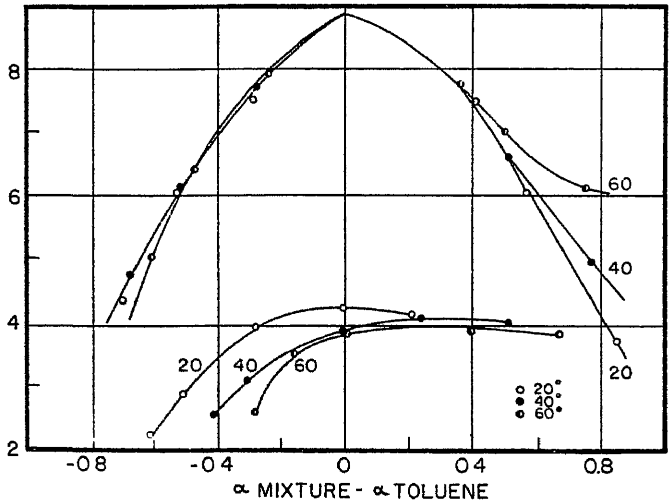


FIG. 7.—Intrinsic viscosity of fraction JNZ-2 as a function of cohesive energy density of mixed solvents.

(B) Precipitation Points as a Function of Cohesive Energy Density.—Another example of the failure of the cohesive energy density concept occurs in the precipitation of polystyrene from toluene, using as precipitant a mixture of *n*-butanol and *n*-heptane. One would expect a range of mixtures of these two nonsolvents to possess an average cohesive energy density near enough to that of polystyrene for the mixture to act as solvent. This is not observed. It is true that addition of small amounts of either precipitant to the other results in a less efficient mixed precipitant. However, the amount of mixed precipitant required to precipitate polystyrene from toluene merely goes through a maximum, and never approaches infinity. This is shown in Fig. 6, where the data are plotted in terms of the average cohesive energy density of the precipitant mixture.

(C) Intrinsic Viscosity as a Function of Cohesive Energy Density.—In Fig. 7 the data of Fig. 1, 2 and 3 are plotted against the interaction constant α rather than against the mixed solvent composition. Curves for the systems toluene-hexane and toluene-butanol fit moderately well the predictions based on cohesive energy densities. Solutions containing methyl ethyl ketone, however, fall completely out of line. Increasing deviations also are observed in the butanol system with increasing amounts of butanol. When asymmetric polar solvents are concerned, therefore, it must be concluded that the intrinsic viscosity is not determined simply by the difference in cohesive energy density between solvent and polymer. Furthermore, in such cases precipitation does not occur at a definite value of average cohesive energy density.

(D) Variation of k' with Solvent Type.—The specific viscosity of a given polymer-solvent system can be expressed as a power series in terms of the concentration of polymer

$$\eta_{sp} = a_1c + a_2c^2 + a_3c^3 + \dots .$$

The first expansion coefficient, a_1, is by definition the intrinsic viscosity $[\eta]$. The second expansion coefficient has been shown by Huggins and others to be proportional to the square of $[\eta]$. Thus, we may write, for small finite concentrations where higher terms can be neglected,

$$\eta_{sp} = [\eta]c + k'[\eta]^2c^2.$$

Here k' is a constant characteristic of a given homologous polymer series in a given solvent. Table I shows the values of k' obtained in different pure and mixed solvents. While these results are scattered, a general trend may be observed. The constant k' tends to be higher in poor solvents and lower in good solvents.

TABLE I

Liquids.	Weight Ratio.	k'.		
		19·9°.	40·4°.	60·4°.
Toluene/Butanol . . .	100/0	0·30	0·30	0·30
	90/10	0·48	0·46	0·49
	80/20	0·44	0·43	0·41
	70/30	0·58	0·47	0·48
Toluene/heptane . . .	86/14	0·42	0·40	0·36
	72/28	0·45	0·43	0·42
	62/38	0·50	0·55	0·50
Methyl ethyl ketone/heptane.	88/12	0·43	0·48	0·55
	76/24	0·46	0·57	0·57
	64/36	0·66	0·54	0·60
	59/41	0·59	0·68	0·74

Polymer Research Institute,
Polytechnic Institute of Brooklyn.

PARTICLE SWELLING AND THE VISCOSITY OF SUSPENSIONS.

By F. Eirich and J. Sverak.*

Received 30th August, 1946.

When theoretical investigations on the viscosity of solutions of high polymers were vigorously advanced beyond the pioneer work of Einstein and Jeffery, it was felt that a parallel experimental attack was needed. Simultaneously with the calculations of Guth and Simha [1] experiments on model suspensions were started in the expectation, firstly, of testing the theory on these well-defined systems more conclusively than on real solutions in which a number of essential features is always a matter of conjecture; secondly, by making model suspensions increasingly similar to real solutions, of facilitating extrapolation towards the constitution of solutions and of learning the relative importance of each step in the process of approximation. The essential steps were: dilute suspensions of perfect spheres, then of rigid rods employing increasing axial ratios; then, by using increasingly smaller particles, to introduce Brownian motion. Further the influence of particle rigidity was to be looked into by making the rods coiled or flexible; finally concentration and mixing effects were to be studied. Circumstances prevented the implementation of the full programme. Only the first two parts were studied in full and reported.[2] Points 3 and 4 of the programme were begun but had to be left incomplete, while the last parts could not be started at all.

We shall now report on the experimental material relating to swollen particles, in which the process of swelling was not so much studied as such, but was used as a means of preparation. The results, therefore, are indicative mainly of the behaviour of soft particles in suspension; they are, however, characteristic of a certain type of swelling. It will also be seen that soft threads may serve in some ways as models of single molecules in their relation to the surrounding solvent.

Experiments and Results.

The same fibres were used, cut from regenerated cellulose (Copper-Rayon), as in the previous experiments. At a constant thickness of 9μ the lengths cut were *ca.* 0·45, 0·90, and 1·8 mm. Winnowing was used as a method of fractionation, removing particle agglomerates. Attempts to soften these fibres, were guided by the following requirements: the fibre length ought to change as little as possible, the boundaries have to remain sharp, the softened fibres must not stick, fuse with each other, or break up easily; furthermore, the softness ought to be such as to allow the particle to bend readily if held at one end and subjected to a rate of shear of the order of 10 sec.$^{-1}$.

---

* The experimental results reported in this paper are largely derived from my former co-worker's, Dr. Sverak's, thesis, which formed part of the fundamental research programme on the viscosity of suspensions and solutions carried out in Prof. H. Mark's Vienna Institute. Publication was delayed by the international situation. Although it is regretted that no contact with Dr. Sverak has as yet been possible, it is felt, in view of his earlier expressed wish to see this paper published, that this should now not be further delayed. Much of the interpretation offered remain, however, my responsibility, and thanks are due to Prof. Mark. for sponsoring this work.

[1] Guth, *Kolloid Z.*, 1936, **74**, 147; 1936, **75**, 15; Guth and Simha, *ibid.*, 1936, **74**, 266; 1936, **76**, 16; *J. Physic. Chem.*, 1940, **44**, 25; *J. Appl. Physics*, 1942, **13**, 147.

[2] Eirich, Bunzl, Macgaratha, Goldschmidt and Simha, *Kolloid Z.*, 1936, **74**, 276; **75**, 20; 1937, **86**, 7; *Akad. Wiss., Wien.*, 1937, **146**, (IIb), 513.

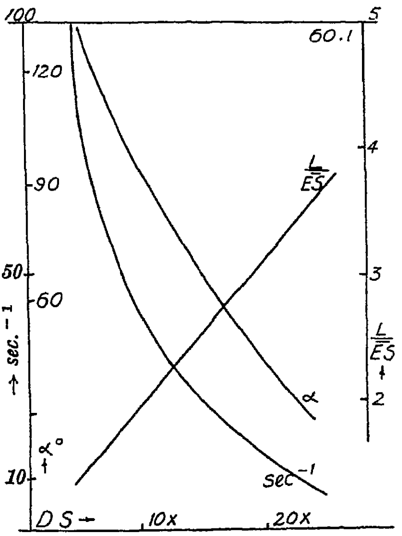


FIG. 5.—Particle curling as function of Degree of Swelling. Fibres of ultimate axial ratio 60 : 1. Softness and curling determined by the rate of shear necessary for fibre bending, by the average angle of kinking, and by the ratio of extended length/average end point separation.

These conditions could not be met by ordinary swelling though a great variety of media was tried. The desired result was obtained by partial xanthation. Our fibres were suspended in 30 % KOH and treated with CS_2 vapour. The cellulose is thereby partially converted into its xanthate, to a degree depending on the duration of the reaction; this time cannot, however, be accurately assessed because of the time required to remove the excess CS_2 by blowing air through the reaction mixture. The highest degree of swelling at which the particles were still safe from disintegration was about 25 times the dry volume and was reached after about 40 min. total action, corresponding to a xanthate formation of about 1 in 4 glucose units. It is remarkable that in this swelling process the fibres do not change their length to any appreciable extent, increasing only in diameter. Softening proceeds parallel with this increase, and to a fairly regular and reproducible extent, so that the square of the ratios of the diameters in the dry and swollen state can be taken as the degree of swelling and may serve as a measure of the softness. In view of the statistical nature of curling, determinations based on measuring individual particles were tried but were too cumbersome. In order to give some idea in absolute terms such as may be used in theoretical treatment, the general dependence on the degree of swelling of rate of shear (bending the particle to a quarter-circle), average angle (between two straight sections of the fibre), and ratio of straight length to end-point separation are plotted further above (Fig. 5).

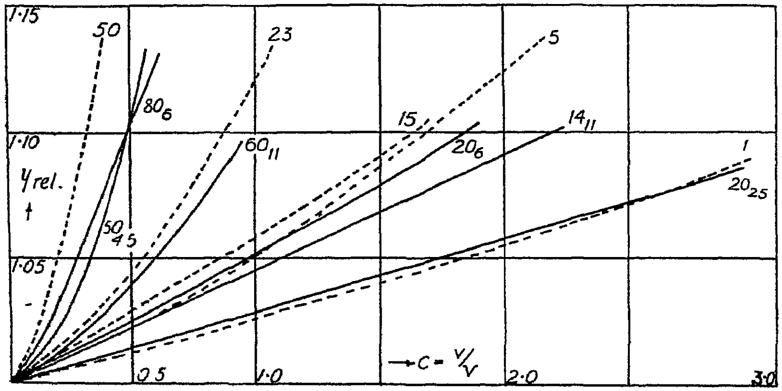


FIG. 6.—Relative viscosity *vs.* concentration.
– – – Rigid fibres. — Swollen fibres. Numbers: axial ratios "extended". Indices: degrees of swelling.

PLATE II

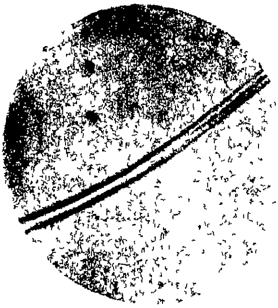


FIG. 1

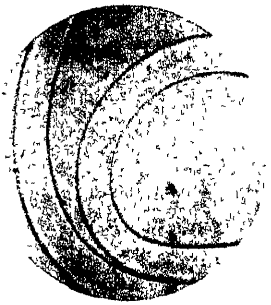


FIG. 2.

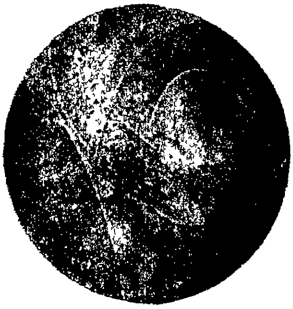


FIG. 3.

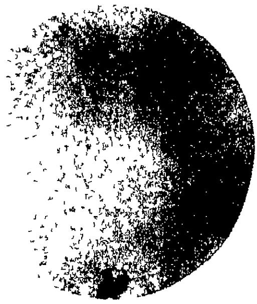


FIG. 4.

FIG. 1.—Dry silk fibre (× 200)

FIG. 2.—Silk fibre, after swelling to 6 times the original volume (× 200).

FIG. 3.—Silk fibre, after swelling to 12 times the original volume (× 200).

FIG. 4.—Silk fibres, 6 times swollen, $L/D = 80$, showing various degrees of curling (× 25).

[*To face page* 59

When the desired degree of swelling was reached and the CS_2 removed, the particles were concentrated and then transferred into strong K-lactate of similar density. This suspending medium proved superior to any other in maintaining the swelling and flexibility of the particles for an unlimited period. A number of particles were then photographed in polarised light and the actual degree of swelling determined. The volume concentration of the suspension follows from the amount weighed in, the density and diameter in the dry state, and the diameter in the swollen state. The latter combined with the (unaltered) particle length gives the axial ratio, L/D, or asymmetry of the extended thread. This ratio, besides their degree of swelling, will be used to classify the particles. Fig. 1 to 4 show respectively a dry rod, the $6x$ and $12x$ swollen state, and a deposit of $6x$ swollen fibres of $L/D = 80$. It was also intended to investigate kinked rigid fibres and to obtain them by de-swelling of our threads. The definition of these particles was, however, poor, and the values of only one example, $11x$, $L/D = 15$, are included.

The suspensions were measured in a double cylinder Couette-apparatus, similar to that already described[3] but considerably improved. Some representative values are given in Table I and the curves in Fig. 6. An attempt was made again to watch the particle behaviour in a glass couette, but in view of their transparency observation was difficult; dyeing was unsuccessful, but it could be observed that there was no noticeable stretching of the coils at the rates of shear employed (10-40 sec.$^{-1}$). Particle orientation, so very marked with rigid fibres, was nearly absent when the coils as a whole were observed, but there was some internal movement.

TABLE I

Degree of Swelling : D.S.(x). Axial Ratio : L/D.
Relative Viscosities : η. Concentration : $100v/V = \%$.

D.S. 4·5–6x.						11x.		10–12x.				25x.	
L/D 20.		50.		80.		15*.		14.		60.		20.	
%.	η.	%.	η.	%.	η.	%.	η.	%.	η.	%.	η.	%.	η.
0·5	026	0·25	030	0·15	026	0·6	035	0·75	034	0·25	017	0·75	020
0·75	040	0·40	070	0·25	050	1·2	070	1·50	068	0·50	040	1·50	042
1·45	076	0·60	145	0·60	128	2·0	130	2·50	115	0·80	077	3·00	085

Discussion.

Our results demonstrate the changes in viscosity following upon softening and curling of suspended fibres. If the particle number is kept constant the loss of asymmetry is as a rule overcompensated by the increase in volume; for constant volume concentration, however, the viscosity drops to a fraction of its former value; the viscosity increase with concentration and with particle asymmetry diminishes greatly.

There is no explicit theory for pliable extended particles without Brownian motion. Taylor's work[4] applies to liquid spheres, Kuhn's[5] and Huggin's calculations[6] to randomly kinked chains in strong Brownian motion. The nearest pertinent approach remains Jeffery's theory,[7] if

---

[*] Fibres partially shrunk and stiffened.

[3] Eirich, Bunzl, Margarethu, *Kolloid Z.*, 1936, **74**, 276.

[4] Taylor, *Proc. Roy. Soc. A*, 1932, **138**, 41.

[5] Kuhn, *Kolloid Z.*, 1933, **62**, 269; 1934, **68**, 2; *Helv. Chim. Acta.*, 1943, **26**, 1394.

[6] Huggins, *J. Physic. Chem.*, 1938, **42**, 911; 1939, **43**, 439.

[7] Jeffery, *Proc. Roy. Soc. A*, 1923, **102**, 163. (Cf. Eisenschitz and Burgers, *2nd. Rep. on Viscosity and Plasticity*, 1938, whose approach, however, leads to less agreement with the experiments.)

effective hydrodynamic ellipsoids are substituted for our curled fibres. Giving specific viscosities of suspensions of rigid rods, too large for B.M., as a function of their axial ratio and orientation, this theory will permit the determination of the asymmetry of the substitute ellipsoids. It must be remembered that the ellipsoids so derived have no other physical meaning (in Kuhn's treatment his " hydrodynamic " dimensions are those of the drag-resistance). In the case of densely coiled, or matted, particles in which much of the solvent is firmly imbibed the substitute ellipsoids will compare more closely with the real shape than in the case of freely draining coils, for which the classical hydrodynamic approach is seriously limited. Simha[8] has recently introduced an exponential relationship between the effective ellipsoid and the degree of polymerisation, without having to make further assumptions about the shape of the molecule in solution, while the statistical model of a randomly kinked coil employed by Kuhn and Huggins can be compared with the effective hydrodynamic dimensions.

In our previous paper agreement between the model experiments and Jeffery's theory was found for the shorter rodlets only. At greater axial ratios our values increased much more steeply, because concentration effects prevented correct extrapolation to zero concentration. The two types of models employed, natural silk and cellulose, moreover, gave two consistent but somewhat different sets of curves ; this can now be explained by the slight swelling of the silk. To deal with particle interaction at very low concentration we may use Guth's[9] extension of Jeffery's theory for higher concentrations wherein he introduces the following approximation for the second term :

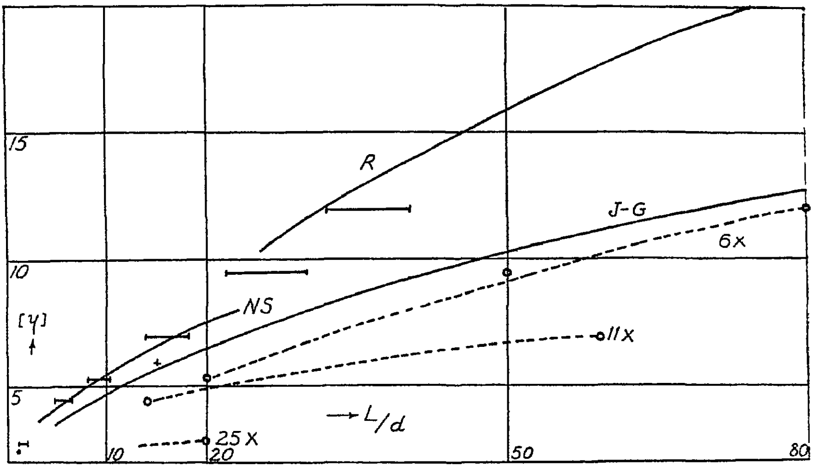


FIG. 7.—Intrinsic viscosities *vs* axial ratios.

—— J-G according to Jeffery-Guth (theoret.).
—— N.S. Model suspensions . . . natural silk.
—— R „ „ . . . rayon.
„ „ . . . swollen.

$$[\eta] = k_1(f + 2) + k_2 L/D f^2 c \; ; \text{ where } f = \frac{L/D}{2 \ln 2 L/D - 3},$$

k_1 is a numerical factor near unity, k_2 a smaller factor depending on the rate of shear, c the volume concentration $= v/V$ and $[\eta] = \eta_{sp}/c_{\to 0}$. The proportionality with about $(L/D)^3$ [as compared with $\sim (L/D)^4$ for B.M.] is slightly too high, but with this approximation the gap between experiment and theory is greatly reduced. The remaining difference is presumably due to rods having to be substituted by appreciably more elongated ellipsoids, and to effects due to the particle size.

---

[8] Simha, *J. Chem. Physics*, 1945, **13**, 188.

[9] Guth and Gold, *Physical Rev.*, 1938, **53**, 322, 926.

Table II contains the recalculated intrinsic viscosities $[\eta]$ of the rigid fibres and the theoretical maximum values, as well as the intrinsic viscosities of the soft-fibre suspensions extrapolated in the same way. Falling very much below the "rigid" values, their corresponding hydrodynamic axial ratios (interpolated from the curve of the rigid rods) are given in column 6, and column 7 shows the ratio between the "extended" and "curled" asymmetry thus derived. For a better illustration the functions of $[\eta]$ against L/D [theoretically rigid rods (silk and cellulose), and soft fibres] are shown in Fig. 7 which demonstrates the method of interpolation and indicates the probable extent of error.

TABLE II

INTRINSIC VISCOSITIES AND AXIAL RATIOS

L/D Rigid.	*L/D* Extended.	$[\eta]_{\text{calc.}}$	$[\eta]_{\text{obs.}}$		*L/D* Coil.	$\frac{L/D\text{ Extended.}}{L/D\text{ Curled.}}$
1 : 1		2·5	2·5			
5		3·6	4·5			
11		5·0	5·5			
	14_{11x}			4·5	5-6	~ 2·6
17		6·0	6·4			
	20_{6x}			5·3	8-10	~ 2
	20_{25x}			2·8	8·1	~ 12
23		7·0	8·0			
32		8·3	12·3			
	$50_{4\cdot5x}$			9·5	22-30	~ 2
50		10·2	16			
	60_{11x}			7·0	14-18	~ 3·7
75		14·3	20			
	80_{8x}			12	32-40	~ 2·4
	15 *			5·8	11-13	

Bearing in mind their length/endpoint separation, or better the ellipsoids which would circumscribe, or contain, the soft particles, the interpolated asymmetries seem surprisingly high. Judging from the microscopical evidence, however, the coils are on the whole so open that the influence of the particles on the liquid inside the loops must be of the same order as on any volume element at the same distance outside. This may further be deduced from the amount of particle interference at rising concentrations as shown by the deviation from linearity. With spheres this interference is first observed at *ca.* 2 %, i.e. at an average distance of the spheres of 4 radii from surface to surface, or 6 from centre to centre, but is not considerable until this distance has shrunk to about one half at 10 %. The same holds approximately for elongated particles which can here be substituted by spheres of appropriate effective radius. These radii (though a function of concentration and rate of shear) remain[10] a small multiple of $\sqrt{ab}$, the ellipsoid's half-axes, approaching *ca.* $3^3\sqrt{ab^2}$

---

* Fibres partially shrunk and stiffened.

[10] *Loc. cit.*, 3 ; also Eirich and Huber (unpublished results) and Eirich and Simha, *Z. Physikal. Chem. A*, 1937, **180**, 447.

at $L/D = 100$; the increase being due to random axial angular distribution which at any average distance between their centres causes frequent closer surface approaches or collisions. On loose back-coiling such close approaches are few compared with the whole length of the chain and those occurring will be in a " T " or " X " fashion rather than alongside, forces excluded. Consequently our coils must be expected to be fairly free draining until they occupy about 10 % of the volume of the enveloping ellipsoid, to be half-draining, i.e. reducing the mobility of the solvent inside the coil to one-half of the mobility outside, at about 20 % density, and to have all internal liquid " immobilised " in a particle matted to 50 % of its dry density (the maximum specific volume in solution will, however, be at lower densities, being a product of hindered volume × degree of hindrance). A similar limit for matted particles was found in model experiments by Eirich, Huber and Mark;[11] though their values refer to the Stokes resistance, there should be little difference to participation in dilatation at high degrees of matting.

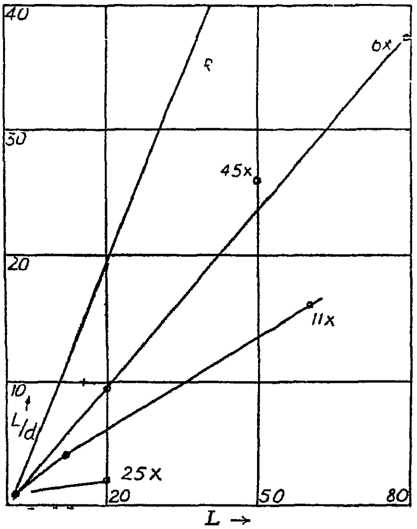


FIG. 8.—Axial ratio *vs.* length.
R . . . rigid fibres. *x* . . . Degree of swelling.

We emphasize that these approximations hold only if all distances concerned are > 10 diameters of the solvent molecules. They cannot easily be extended to molecular chains of a cross-section about equal to that of the solvent molecule. There are, however, certain indications. Kuhn, e.g., has shown that obedience to Staudinger's rule indicates a freely draining coil, while proportionality with M to a power less than 1 indicates increasing immobilisation. He prepared large models of several degrees of coiling. The less entangled ones, standing for medium molecular weights and free draining, are strikingly similar to our more asymmetric models, see Fig. 4 (Plate II). Hence it appears that our fibres may be quite comparable although it may be assumed that on a molecular scale, notwithstanding a correction for slip, the same degree of coiling ought to produce as a rule a higher degree of hindrance.

From this, as well as from microscopic evidence (e.g. small objects in full B.M. inside the loops) our particles are certainly free draining Being free draining, most of their length is hydrodynamically active. The key as to what theory to apply lies now in the particle orientation. If there were full B.M. (i.e. random orientation), Kuhn's procedure would be applicable, and $[\eta]$ ought to be proportional to the fibre length. This is clearly not the case; $[\eta]$ increases less than $\sim L$, in fact very much as given by Jeffery's theory if plotted against the extended length times a factor. This is brought out very clearly by the effective axial ratios interpolated as above (though not from the theoretical but from the experimental curve for rods) (Fig. 8). There is a clear functional relation between these effective ratios and particle length being linear for the lower degree and very nearly so for the higher degree of swelling. The above adopted procedure of substituting rigid effective ellipsoids for our coils is thus the more justified the lower the degree of swelling.

---

[11] Eirich, Huber and Mark, *Papierfabricant*, 1937, 35, 251.

The inference is that for constant swelling the curvature introduced is the same throughout, i.e. a "waviness" that repeats itself along the fibre. It was not realised before that the effective axial ratio of coiled fibres may be proportional to their full length; that is to say, that the statistical outer dimensions are not the hydrodynamically effective ones, at least not for loose, free draining coils without B.M. As regards the proportionality constant, Fig. 9 shows that the first curvature introduced influences the viscosity relatively more than later, more extensive, coiling. At the same time the proportionality causes a much greater drop in relative viscosities with the more asymmetric particles. The fitting into the viscosity function for rods indicates that the particle orientation of our curls is, in its effects, not greatly different. Long rods lie mainly parallel to the streamlines; the coils so far as could be observed show no conspicuous orientation, but it may be assumed that within each coil the mobility of the segments allows many of them to rotate readily into a position more or less parallel to the streamlines. As will be seen later the existence of a fair internal mobility is supported by the concentration dependence. If this be so, the clue to the relation between effective and real shape is that the interpolated axial ratios are those of the mobile segments rather than of the enveloping ellipsoid. In line with this, Kramers[12] in the course of a new treatment has shown that a chain, closed to a ring raises the viscosity half as much as when open. Although this applies to the case of Brownian movement, it agrees with our findings when this is allowed for. Where the internal movement is hindered, as in the shrunk particles 15 : 1, the viscosity rises at once nearly to that of a rigid rod.

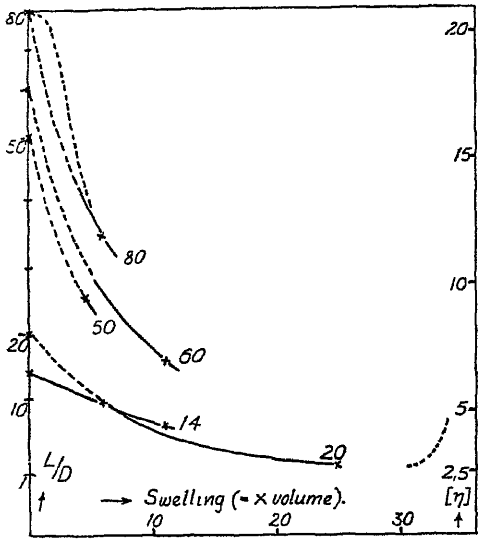


FIG. 9—Intrinsic viscosities *vs.* degree of swelling. ("Extended" axial ratio constant).

The viscosity reduction on softening becomes even more striking if, instead of the intrinsic viscosities, concentration effects are investigated. The majority of the curves in Fig. 6 lie below those of rigid rods of the same asymmetry, and none rises above. The reason for this must again be sought in the internal mobility. In view of the stability of total length and cross-section it is unlikely that there is much flow within the body of the particles, the effect considered by Taylor. The continuous surface rotation in the velocity gradient deforming the coil within its average shape is much akin, however, to such an internal flow, since it involves periodic displacements, both inside the thread and in the liquid between its segments. It is this somewhat emulsoid character of our suspensions which is probably an additional factor in reducing the effective axial ratios and becomes of greater importance during particle interaction. The additional motion caused in the solvent by the presence of one particle need not drop to zero at the surface of a neighbouring internally mobile coil (and be compensated by an opposite motion) but can continue into it and be there partially dissipated. It is as if the particles were more permeable and less responsive to the disturbance in the solvent set up by their own presence. The effect can be seen in its extreme with the very soft $25x$ swollen fibres whose effective shape is not only reduced to that of a sphere, but where the viscosity rise with concentration may fall below that of rigid spheres. Molecular chains may in this respect be less soft than our models.

---

[12] Kramers, *J. Chem Physics*, 1946, 14, 415.

Although the extreme swelling of cellulose during xanthation is well known it would be of interest to consider the mechanism of the observed preservation of shape at such high degrees of swelling. It appears to be established[13] that the alkali enters the fibres with the water during the wetting, and forms inter- and intra-micellar cellulosate. This sets up an osmotic pressure and reduces the cohesion, facilitating further swelling, until equilibrium between the osmotic and the internal pressure is reached. In our case about 200 % swelling ($3x$) in 5·5 KOH occurs. This permits the CS_2 to enter and to form at random a certain amount of xanthate, predominantly in the amorphous region. The crystallite fibrils are largely preserved, though some surface action will increase the amorphous portion at the expense of the fibrous elements. The xanthate groups act as wedges, opening the amorphous network further up till at about 2 glucose residues to 1 xanthate, dissolution occurs.

From our own microscopic observations we suggest that the remaining fibrils, preserving a skeleton of the old oriented structure, fulfil the function of a filler, acting as a longitudinal vulcanisation. The swelling body will thereby retain a great deal of the original shape and resist deformation. Swelling will be largely radial: there is little resistance to pushing the fibrils apart but much friction and obstruction if the gel seeks to expand (or contract) along the fibrils, i.e. the amorphous mass will flow much faster radially than axially. The essential feature is therefore the presence of non-extensible longitudinal elements and the absence of cross-links between them; this explains the observed degree and anistropy of swelling and is compatible with the body of other knowledge. There should be obvious applications for this suggested mechanism in many cases of swelling of other high polymers.

Summary.

Suspensions of soft, curled, fibres were prepared and the viscosities measured in a Couette apparatus. The dimensions and the degree of curling of the fibres were determined by microscope and their behaviour observed in a glass Couette.

The intrinsic viscosities were compared with those of rigid fibres, of the same (extended) axial ratio, as determined by earlier model experiments. The rigid fibres had been found to be in functional but not in quantitative agreement with Jeffery's theory. For soft and curled fibres all values lie consistently lower. By interpolating them on the curve for rigid fibres effective hydrodynamic ellipsoids can be ascribed to the soft fibres.

The axial ratios of these effective ellipsoids increase linearly or nearly so with the extended length of the fibres. However, these asymmetries are much greater than of ellipsoids enveloping the curls as observed in the microscope. It follows that for soft fibres too, the law connecting intrinsic viscosities with particle asymmetry must be closely related to that derived by Jeffery, and that the curling reduces the effective hydrodynamic length in a linear way. This in turn indicates that the particles are " free draining ", in agreement with the microscopic evidence, and also with considerations by Kuhn for similar particles. The concept of free draining is then discussed in general terms of effects of concentration in suspensions.

---

[13] Ott, *Cellulose* (Interscience, 1943); cf. article by Kline, Sisson, Nicoll and others.

The increase of viscosity with concentration for the curled particles is remarkably small, the smaller the softer the fibres. It follows that not only the magnitude, but also the " reach " of the disturbance caused in a liquid by soft particles is reduced.

Finally, the extent and the anisotropy of regenerated fibre swelling is discussed in terms of the current views of swelling reactions, and explained by the effect of the persisting crystalline fibrils.

Dept. of Colloid Chemistry,
Cambridge.

GENERAL DISCUSSION *

Dr. J. J. Hermans (*Utrecht*) (*communicated*) : In connection with the interesting experiments dealt with by Dr. Eirich, it may be of interest to remember that a molecule which in sedimentation experiments behaves like a " matted " particle, where the solvent is immobilised to a considerable extent, may show almost complete free draining in viscosity determinations. This is due to the fact that the immobilisation of the liquid is primarily a purely hydrodynamic effect : the liquid moving in any volume element drags away a certain amount of liquid in neighbouring volume elements by a Stokes-Navier flow. Clearly, this effect must be much more pronounced in sedimentation, where all parts of the molecule move in the same direction, than in viscosity experiments (i.e. in a velocity gradient) where the velocity of one-half of the molecule with respect to the solvent is opposite to that of the other half of the molecule.

Mr. V. R. Gray (*Manchester*) (*communicated*) : For dilute polymer solutions which obey the equation

$$\eta_{sp} = [\eta]c + k'[\eta]^2c^2$$

the quantity k' is an important measure of cohesional forces between polymer molecules, or between different parts of the same molecule. This is clearly shown in Huggins' [1] derivation. The higher value of k' obtained for solutions in poor solvents is thus due to the increased tendency of the molecules to aggregate.

It was found that dilute solutions of aluminium soaps in benzene obeyed the above equation, but that the value found for k' was very much higher [2] than that found for polymers, namely, 10·2. Since the viscosity at infinite dilution was of the order expected for a dispersion of single molecules or small aggregates it was concluded that the cohesional forces between the soap molecules are very strong, approaching those causing chemical bond formation. In this way more concentrated solutions of aluminium soaps possess all the properties of polymer solutions of high molecular weight.

Dr. E. M. Frith (*Cambridge*) said : Measurements of the viscosity relationships of synthetic rubbers, polyvinyl acetate and polyvinyl chloride in mixed solvents and plasticisers [3] showed wider variations in the slopes of the $\eta_{sp}/c - c$ curves than in the intrinsic viscosities themselves. Huggins' k' values are not useful here as $[\eta]$ may remain nearly constant for a given polymer in a series of solvent mixtures while the initial slope of the curve varies considerably. While molecular uncoiling in better solvents would undoubtedy lead to greater chain-chain interaction and thus to high viscosity slopes, the wide differences found experimentally may also be due to *inter*molecular structural build-ups in the different solvents. Different types of interaction between chain-segments could lead to temporary aggregations at finite concentrations without affecting the molecular configurations in the region of infinite dilution where the chains move independently. If molecular aggregations occur preferentially by end-to-end aggregations of chains, both statistical uncoiling and aggregation will favour high viscosity slopes but in other types of system, different types of aggregation may well oppose the configurational tendency and lower the overall solvent effect. High viscosity slopes in the systems studied are related to high swellings in the poorer component of the solvent mixture and similar molecular interactions may determine the behaviour of the swollen polymers. It seems unlikely that an exact description of the viscosity curve can be given in thermodynamic or statistical terms alone.

---

* On two preceding papers.

[1] Huggins, *J. Amer. Chem. Soc.*, 1942, 64, 2716.

[2] Gray (forthcoming publication).

[3] Frith, *Trans. Faraday Soc.*, 1945, 41, 90.

THE DETERMINATION OF MACROMOLECULAR CONFIGURATIONS IN DILUTE SOLUTION BY LIGHT SCATTERING.*

By P. M. Doty,† W. A. Affens and B. H. Zimm.‡

Received 26th July, 1946.

Introduction.

When light is scattered from a dissolved molecule the size of which exceeds a few hundred Ångstroms, the angular dependence of the scattered light is not symmetrical about the transverse direction. The relation of this effect to a characteristic dimension of the molecule has been calculated for spherical, rod-like and randomly coiled particles in infinitely dilute solution.[1, 2, 3] A portion of the present study serves to test these relations for rod-like and coiled molecules. In addition, this investigation consists of measurements of the ratio of the intensities of the forward to the backward scattered light (dissymmetry) for solutions of polystyrene fractions at various concentrations in different solvents. Supplementary measurements of viscosity and osmotic pressure are included. A few measurements on tobacco mosaic virus, which demonstrate the correctness of the method, are also reported.

These measurements may be interpreted to show the effect of solvent ability (solvent-solute interaction) on the configurations of polystyrene molecules: this effect is compared with the variation of intrinsic viscosity in the same solvents. It is also demonstrated that the dissymmetry may be greatly concentration dependent. An explanation of this behaviour is offered. Finally the change of dissymmetry with the molecular weight of a series of polystyrene fractions provides some additional information on the configurations of polystyrene molecules in solution. The general correctness of the optical technique has been verified by its application to solutions of tabacco mosaic virus.

Theoretical Considerations.

The theory and practice of the determination of molecular weights of large molecules by observing the intensity of the light scattered from their solutions have been described previously.[1, 4, 5] The determination of dimensions has also been described,[2, 3, 5] but for convenience we will summarise the theory here.

---

* Taken (in part) from a thesis presented in partial fulfilment for the degree of M.Sc., Polytechnic Inst. of Brooklyn, 1946.

† Present address: Department of Colloid Science, Cambridge University, Cambridge, England.

‡ Present address: Department of Chemistry, University of California, Berkeley, California, U.S.A.

[1] Debye, P. P., *Thesis* (Cornell University, 1944); Rubber Reserve Report CR110; *J. Appl. Physics*, 1946, 17, 392.

[2] Debye, P., Lecture, Polytechnic Institute of Brooklyn, W.Y., Nov. 25, 1944; *J. Physic. Chem.* (in press).

[3] Zimm, Stein and Doty, *Polymer Bull.*, 1945, 1, 90.

The method is based on the observation of the *intra*molecular interference that occurs when the scattering particle is comparable with the wave length of the light. The observation consists of measuring the ratio of the relative intensities of the light scattered at two different angles from a solution of the particles. For a particular model particle it is possible to calculate the value of this ratio as a function of the angles of observation and the size of the particle. These calculations have been made for thin rigid rods with random orientation and for freely coiling chains. (A freely coiling chain is defined as one in which the probability, W(r), of finding a distance r separating any specified pair of links of the chain is given by a Gaussian expression :

$$W(r) = 4\pi r^2\left(\frac{3}{2\pi R^2}\right)^{3/2} e^{-3r^2/2R^2} \quad . \quad . \quad . \quad (1)$$

where R is the root mean square distance between the pair of links.)

The expression for the ratio of the intensities I_1 and I_2 measured at angles θ_1 and θ_2 (the angles being measured backward from the direction of scattering) is the following for a thin rod :

$$\frac{I_1}{I_2} = \frac{(1 + \cos^2\theta_1)\left[\frac{1}{x_1}\int_0^{2x_1}\frac{\sin u}{u}\,du - \left(\frac{\sin x_1}{x_1}\right)^2\right]}{(1 + \cos^2\theta_2)\left[\frac{1}{x_2}\int_0^{2x_2}\frac{\sin u}{u}\,du - \left(\frac{\sin x_2}{x_2}\right)^2\right]} \qquad (2)$$

where $x = 2\pi\frac{L}{\lambda}\sin\frac{\theta}{2}$, L is the length of the rod and λ is the wave length of the light in solution.

The corresponding expression for the coil, involving R, the root mean square end-to-end separation of the chain, follows :

$$\frac{I_1}{I_2} = \frac{(1 + \cos^2\theta_1)v_2^2\,[e^{-v_1} - (1 - v_1)]}{(1 + \cos^2\theta_2)v_1^2\,[e^{-v_2} - (1 - v_2)]} \quad . \quad . \quad (3)$$

where

$$v = \frac{8}{3}\left(\frac{\pi R\sin\theta/2}{\lambda}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad (4)$$

A similar expression has been obtained for spherical particles but this is omitted here because we have no measurements available on spherical particles.

If we define [1] a *dissymmetry coefficient*, q, as one less than the ratio of the intensities scattered at a given pair of angles symmetrical about the transverse direction ; that is, $q = (I_1/I_2) - 1$ where $\theta_3 - 90° = 90° - \theta_1$. This coefficient can be calculated either as a function L or R. Plots of this coefficient are shown in Fig. 1. Since the angle of observation in our apparatus depends on the refractive index of the solution, plots are given for several typical solvents. These actually correspond to slightly different angles of observation. The angles used in calculating the plot for the rod-like particle are those obtaining when water is the solvent.

It is thus possible by measuring q to determine either L or R. It should be noted that the calculations are exact and quite free from ambiguity as long as (1) the refractive index difference between the particle and the solvent is not too large ; (2) the particles actually conform to the models assumed for them ; and (3) the particles scatter light independently of one another. Condition (1), which arises from the possible distortion of the incident light by the scattering particles, is felt to be negligible on physical grounds for the systems observed here. Like-

---

[4] Doty, Zimm and Mark, *J. Chem. Physics*, 1944, **12**, 144 ; 1945, **13**, 159.

[5] Stein and Doty, *J. Amer. Chem. Soc.*, 1946, **68**, 159

wise condition (2) is probably satisfied. Condition (3), however, must be given serious attention.

If condition (3) is not met, the dissymmetry coefficient will contain a contribution from *inter*molecular scattering interference as well as from the *intra*molecular interference that we wish to observe. It proved to be impossible in practice to go to low enough concentrations to eliminate the intermolecular effect, except in specially chosen solvents. Consequently, it was necessary to devise a means of extrapolation. Fortunately theory predicts what the concentration dependence of the dissymmetry coefficient will be: where there are negative deviations from Raoult's law (the usual case in polymer solutions) the dissymmetry coefficient will decrease with increasing concentration; where there are positive deviations from Raoult's law it will increase, and where there are no deviations there will be little concentration dependence and the intermolecular effect will be negligible. The former case corresponds to " good " solvents and the latter two to " poor " solvents. In any case the concentration dependence of the dissymmetry coefficient will be linear at sufficiently low concentrations.

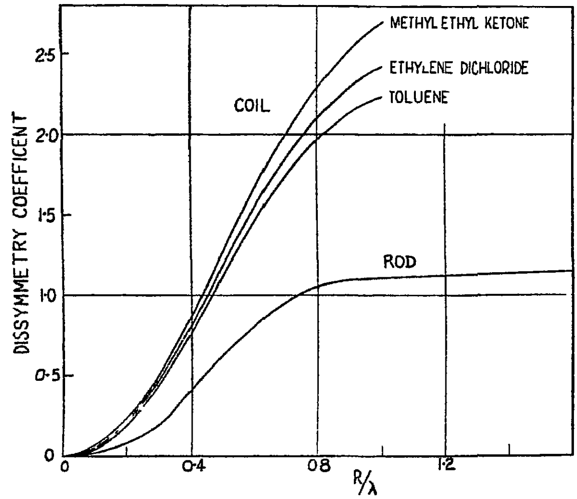


FIG. 1.—The relation between the dissymmetry coefficient and the root mean square separation of the ends of a chain-like molecule or the length of a rod-like molecule.

These relations, whose qualitative results have been stated, follow from the close connection between the Fourier transform of the pair distribution function of the solute particles in solution, which determines the dissymmetry coefficient, and the " cluster integral " involving the same distribution function, which determines the extent of the deviations from Raoult's law.[3, 6]

Therefore an extrapolation of the dissymmetry coefficient of the solute particles to infinite dilution should yield a reliable result if the measurements extend into the low concentration region where the dependence is linear. The limiting value of q as the concentration approaches zero is a molecular constant, which we will designate by $[q]$, the ***intrinsic dissymmetry coefficient***. From $[q]$ one may calculate the R or L values of the dissolved particle if so desired. The value of $[q]$ is a weighted measure of the extension in space of the matter composing the molecule and is, of course, an independent quantity whose interpretation in terms of a model is optional just as in the case of intrinsic viscosity.

---

[6] Zimm, *J. Chem. Physics*, 1945, 13, 141; 1946, 14, 164.

Experimental.

Apparatus.—The apparatus used in this work was identical with that described by Stein and Doty.[5] The dissymmetry coefficient was obtained from the measurement of the relative intensities of light scattered at two angles symmetrically placed about a perpendicular to the incident beam. These angles vary somewhat from solvent to solvent, but were approximately 45° and 135°. The apparatus was checked for symmetry of construction by measuring the fluorescence of a very dilute solution of fluorescein, which gave equal intensities at all pairs of angles of observation. The light used in all measurements was the 5460 A. line of a mercury lamp.

For several purposes to be described later it was necessary to have the intensity of scattering from the various solutions all measured against the same arbitrary standard. The apparatus for this purpose, which was also described previously,[5] measured intensities at an angle of 90° to the incident beam. The arbitrary standard used was a suspension of zinc oxide pigment in a highly viscous solution of polystyrene in acetyl tributyl citrate, which was chosen for its clarity and low volatility.

The osometers used in this work were the semi-micro type described by Zimm and Myerson[7] Denitrated collodion membranes were employed.

Preparation of Polystyrene Samples.—Fraction RT-H was prepared in the following manner: the monomer was washed with alkali to remove inhibitors, distilled at 15 mm., then allowed to stand six months at room temperature. The polymer was precipitated by pouring the partially polymerised (about one third) mixture into alcohol. A broad fraction was obtained by then precipitating about one-fourth of the sample from a 1 % solution in butanone with methanol.

Fractions JNZ-1 and 4, and JN-7 were prepared from a different sample of polystyrene obtained in a manner similar to that described for fraction RT-H. From this sample fractions were prepared by two successive re-precipitations.

Fraction BZO-4 was one of several prepared in a way similar to the JNZ fractions from a polymer sample which had been obtained from a room temperature polymerisation in which 0·2 % benzoyl peroxide had been used as catalyst.

Fraction RT-H which must have had a broad molecular weight distribution, was used for the study of solvent variation exclusively. Fraction BZO-4 was probably much sharper and was used for the same purpose. The other fractions were carefully prepared so as to obtain as narrow molecular weight distributions as possible for their use in molecular weight and size investigations.

Solvents and Solutions.—In order to cover a range of solvent power and yet in each case maintain a sufficiently large refractive index difference necessary for high turbidity, the following solvent media were chosen: (in order of decreasing solvent power or ability) toluene, ethylene dichloride, butanone and butanone-*iso*propanol mixture (85 : 15 by volume). The two components of the mixture have nearly equal indices and densities so that selective adsorption is not reflected in the observations.

All solutions were filtered under pressure through Corning " fine " sintered glass filters (about 15 micron pore size) to render them optically clear; the JNZ solutions were also filtered through a fine asbestos filter to remove haze which persisted after passage through the glass filter.

Dissymmetry and turbidity measurements were made on all solutions at successive dilutions down to 0·0313 g./100 ml., each solution at each concentration being filtered into the cell through the glass filter. The repeated filtrations were necessary, since particles of foreign matter of sizes in the range of the dissolved molecules would contribute greatly to the observed dissymmetry. It was believed that the solutions investigated were rendered satisfactorily clear through the repeated filtrations.

Viscosity.—The viscosities of the dilute solutions were measured in Ostwald viscometers. The flow times for the solvents were always more than 100 seconds and hence the kinetic energy corrections were negligible. The specific viscosity was calculated from the formula $\eta_{sp} = (t - t_0)/t_0$; t is the time of flow of the solution and t_0 the time of flow of the solvent. The intrinsic viscosity was taken as the limit at zero concentration of η_{sp}/c, where c is the concentration in grams per 100 ml. of solution.

Molecular Weights by Light Scattering.—The molecular weights of four polystyrene fractions were determined from measurements of the absolute turbidity and the dissymmetry in the manner described, for example, by Stein and Doty.[5] In this case the calibration of the turbidimeter was based on the scattering from a solution of low molecular weight polystyrene in butanone which exhibited negligible dissymmetry. The absolute turbidity of this solution had been determined by transmission at a number of wavelengths by means of a Beckmann spectrophotometer. By plotting the observed extinction coefficient (turbidity) against the logarithm of the wavelength a straight line was obtained from which the turbidity at 5460 A. could be read with enhanced precision compared with that of a single direct measurement.

---

[7] Zimm and Myerson, *J. Amer. Chem. Soc.*, 1946, 68.

Solvent Correction to Dissymmetry Coefficient.—The theoretical formulae for the dissymmetry coefficient are to be applied to the light scattered from the solute only. However, the observed scattering comes from both the solvent and solute. The angular dependence of the light scattered from the solvent, which is composed of small molecules, is given by the expression $(1 + \cos^2 \theta)$. Hence the observed dissymmetries must be corrected for the presence of light scattered from the solvent.

Since we can measure the relative intensities of solution and solvent only at 90° by means of the turbidimeter, it is necessary to estimate the relative intensities of the solution and solvent scattering at the two angles of observation of the dissymmetry apparatus. From the estimated intensities, the dissymmetry coefficient of the scattering due to the solute alone can be determined. As a result a correction formula for the observed dissymmetries is obtained.

The following symbols are used: θ_1 and θ_2 for the two angles of observation, which in our instrument are chosen to be equally distant from the incident beam ($\theta_2 > \theta_1$); the intensities of the "solvent" scattering at these angles and 90° are I_1°, I_2° and I_{90}°; the corresponding intensities for the solute, I_1', I_2', I_{90}'; the corresponding intensities for the solution I_1°, I_2° and I_{90}°; the solution dissymmetry coefficient, q'; the corrected dissymmetry coefficient (of the solute), $q = (I_1 - I_2)/I_2$; finally let $r = I_{90}^\circ/I_{90}'$, the ratio of solvent to solution intensities at 90°. Since θ_1 and θ_2 are symmetrically placed about 90°, $\cos \theta_1 = \cos \theta_2$ and $I_1^\circ = I_2^\circ$. Since the cosine of 90° is zero, $I_1^\circ = I_2^\circ = rI_{90}' (1 + \cos^2 \theta)$. Figure 2 is a plot which illustrates the angular variation of the intensity divided by $(1 + \cos^2 \theta)$

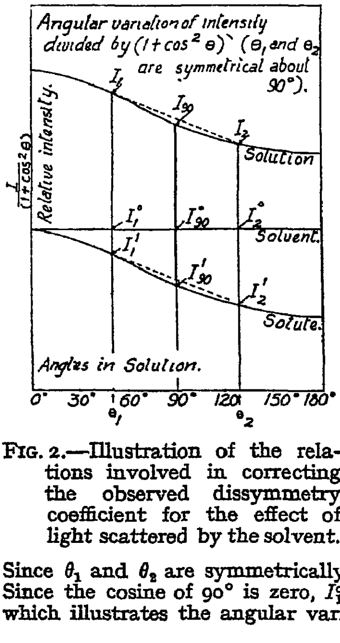


FIG. 2.—Illustration of the relations involved in correcting the observed dissymmetry coefficient for the effect of light scattered by the solvent.

TABLE I.—OBSERVED DISSYMMETRY COEFFICIENTS OF POLYSTYRENE FRACTIONS.

Fraction.	Solvent.	Concentration, g./100 ml.					
		1·000.	0·500.	0·250.	0·125.	0·0625.	0·0313.
RT-H	Ethylene dichloride	0·20	0·24	0·43	0·75	0·90	1·05
,,	Toluene	—	0·42	0·45	0·66	0·78	0·88
,,	Butanone	0·43	0·46	0·71	0·99	1·13	1·19
,,	Butanone*iso*-propanol	1·00	1·00	1·13	1·16	1·13	1·12
BZO-4	Ethylene dichloride	—	0·11	0·16	0·21	0·31	—
,,	Butanone	—	0·16	0·29	0·35	0·42	0·46
,,	Butanone-*iso*propanol	—	0·28	0·27	0·30	0·34	—
JNZ-7	Butanone	—	0·23	0·33	0·43	0·52	0·56
JNZ-4	,,	—	—	0·56	0·79	1·02	1·13
JNZ-1	,,	—	—	0·67	1·04	1·38	1·54

for solution, solvent and solute for the case of a freely coiling chain with a dissymmetry coefficient of 0·6. It can be seen that the portion of the curve between θ_1 and θ_2 is approximately a straight line. Therefore, we shall assume that $I'_{90}/(1 + \cos^2 90)$ is the arithmetic mean of $I'_1/(1 + \cos^2 \theta_1)$ and $I'_2/(1 + \cos^2 \theta_2)$. Manipulation of these relations leads to expressions for the intensity of the solute scattering at the two dissymmetry angles. These are, $I_1 = I'_1 - r(I'_1 + I'_2)/2$; $I_2 = I'_2 - r(I'_1 + I'_2)/2$. Therefore

TABLE II.—INTENSITY OF TRANSVERSELY SCATTERED LIGHT IN ARBITRARY UNITS FOR SOLUTIONS OF POLYSTYRENE FRACTIONS.

Fraction.	Solvent.	Solvent Scat.	Concentration, g./100 ml.					
			1·000.	0·500.	0·250.	0·125.	0·0625.	0·0313.
RT-H	Toluene	2·8	33·6	31·3	32·7	25·2	18·0	13·1
,,	Ethylene dichloride	1·5	46·2	53·4	48·0	38·2	26·1	14·0
,,	Butanone	1·2	358	315	236	158	93·6	48·4
,,	Butanone-*iso*-propanol	1·5	3231	1411	678	336	161	80·9
BZO-4	Ethylene dichloride	1·3	70·5	60·8	44·1	27·4	—	—
	Butanone	1·1	304	195	130·5	67·5	40·0	—
	Butanone-*iso*-propanol	1·6	675	327	165	87	—	—
JNZ-7	Butanone	1·5	—	207	149	95·9	53·5	30·1
JNZ-4	,,	0·6	—	—	202	152	84·6	45·5
JNZ-1	,,	0·6	—	—	234	172	105	54·6

$$q = \frac{q'}{1 - r/2\,(q' + 2)} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

With this formula it is possible to calculate the dissymmetry coefficient q of the solute particles from the measured dissymmetry coefficient, q'.

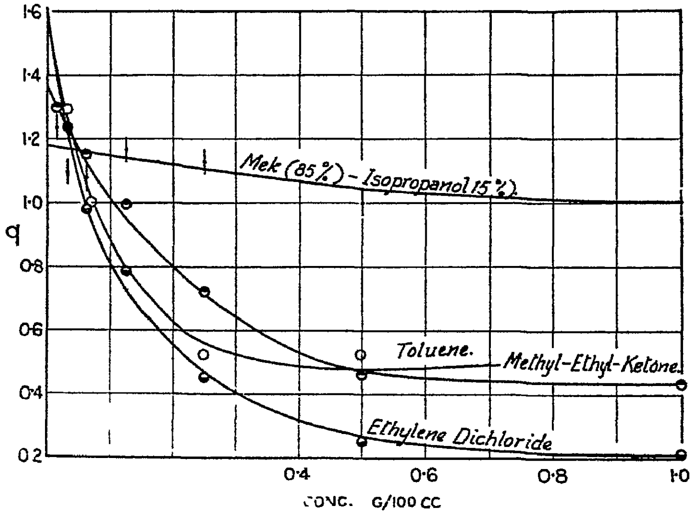


FIG. 3.—The variation of the dissymmetry coefficient with concentration for polystyrene (fraction RT-H) in several solvents.

Experimental Results.

Dissymmetry Coefficients of Polystyrene.—The observed values of q' are given in Table I. In Table II there are recorded the observed values of the intensity of the transverse scattered light relative to an arbitrary standard. By means of equation (5) the observed dissymmetry coefficients are corrected for the effect of scattering from the solvent. The corrected dissymmetry coefficients, q, are given in Table III and are plotted against concentration, c, in Fig. 3 and 4.

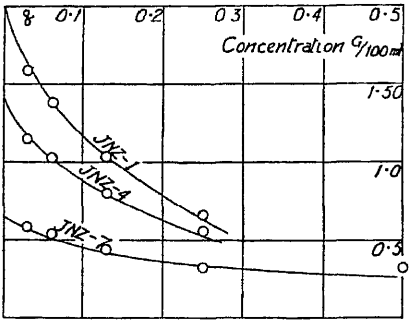


FIG. 4.—The variation of the dissymmetry coefficient with concentration for several polystyrene fractions in butanone.

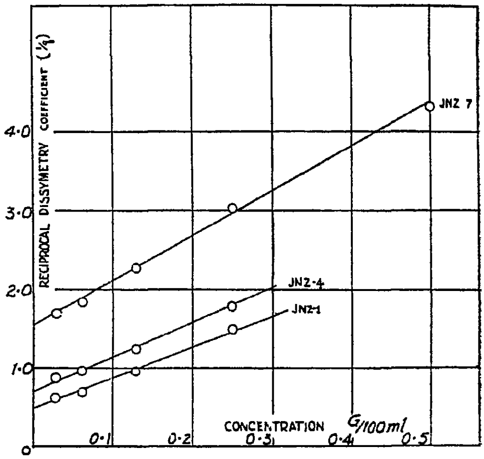


FIG. 5.—Plot of the reciprocal of the dissymmetry coefficient against the concentration for polystyrene fractions of different molecular weight.

TABLE III.—CORRECTED DISSYMMETRY COEFFICIENTS AND ROOT MEAN SQUARE CHAIN LENGTHS FOR POLYSTYRENE FRACTIONS.

Fraction.	Solvent.	Concentration, g./100 ml.							
		1·000.	0·500.	0·250.	0·125.	0·0625.	0·0313.	$c=0$.	R (A.).
RT-H	Toluene	—	0·47	0·50	0·78	1·00	1·28	1·59	2370
,,	Ethylene dichloride	0·21	0·25	0·45	0·79	0·98	1·25	1·64	2390
,,	Butanone	0·43	0·46	0·72	1·00	1·15	1·24	1·39	2100
,,	Butanone-*iso*propanol	1·00	1·00	1·13	1·17	1 14	1·15	1·19	1900
BZO-4	Ethylene dichloride	—	0·11	0·17	0·22	0·34	—	0·45	1100
,,	Butanone	—	0·16	0·29	0·35	0·43	0·52	0·53	1230
,,	Butanone *iso*propanol	—	0·28	0·27,	0·31	0·38	—	0·35	990
JNZ-7	Butanone	—	0·23	0·33	0·44	0·54	0·59	0·64	1350
JNZ-4	,,	—	—	0·56	0·80	1·03	1·15	1·41	2100
JNZ-1	,,	—	—	0·67	1·04	1·39	1·59	2·00	2770

It was found that the extrapolation to infinite dilution to obtain $[q]$ was greatly facilitated if $1/q$ were plotted against c. The resulting plots appear to be linear below a concentration of 0·5 g./100 ml. Such plots are shown in Fig. 5. The limiting value of q, that is $[q]$, and the values of R calculated from $[q]$ assuming that the molecule is a randomly coiled, linear chain are also given in Table III.

Molecular Weight Data.—The osmotic pressure data are shown in Fig. 6 where the osmotic pressure divided by RTc is plotted against concentration.

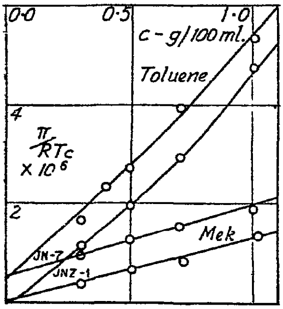


FIG. 6.—Osmotic pressure data for fractions of polystyrene in toluene and butanone.

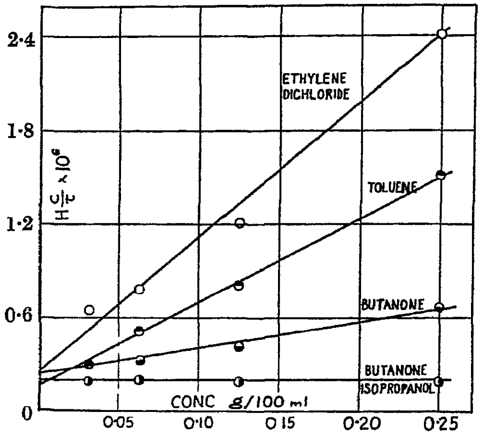


FIG. 7.—Plot of light scattering data for polystyrene fraction RT-H in several solvents.

Each point on the graph is the average of at least two measurements in different osmometers. The intercept of the plot is equal to the reciprocal of the number average molecular weight. The intercepts for fraction JNZ-7 corresponds to a molecular weight of 2,000,000 and that for JNZ-1 is estimated to be of the order of magnitude of 10,000,000.

The absolute turbidity as a function of concentration was measured for fraction RT-H in the four solvents and for the three JNZ fractions in butanone.

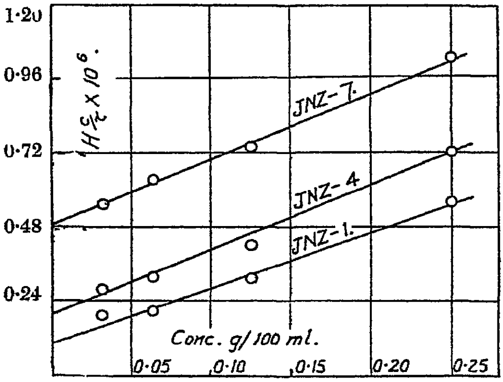


FIG. 8.—Plot of light scattering data for several polystyrene fractions in butanone.

After correction by means of the dissymmetry coefficient for intramolecular interference the customary plots of Hc/τ against c were plotted (see Fig. 7 and 8). The reciprocal of the intercept is the weight average molecular weight. For fraction RT-H in four solvents this value is 4,000,000 ± 400,000. For fractions JNZ-1, 4 and 7 the molecular weight appears to be 9,800,000, 4,800,000 and 2,100,000 respectively.

Viscosity Data.—The viscosity data are shown in Fig. 9 and 10 where η_{sp}/c is plotted against c. The intrinsic viscosities of the JNZ fractions were supplied by Alfrey, Justice and Nelson. The viscosity data recorded was obtained at 26°, however, measurements at 60° showed that changes in $[\eta]$ of less than 10 % in all solvents from the lower to the higher temperature

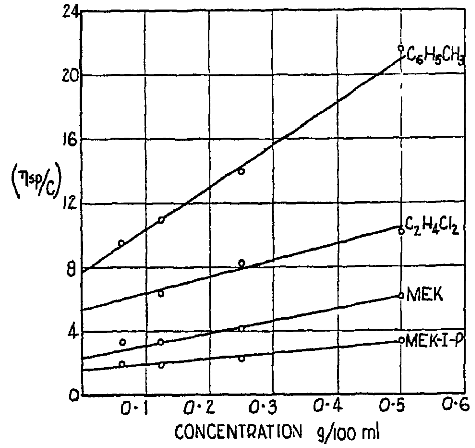


FIG. 9.—Plot of viscosity data for polystyrene fraction RT-H in several solvents at 26° c.

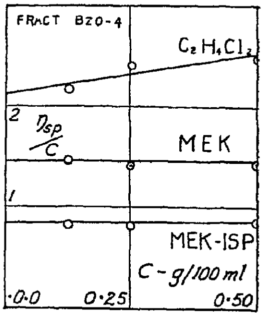


FIG. 10.—Plot of viscosity data for polystyrene fraction BZO-4 in several solvents at 26° C.

Discussion.

The foregoing results demonstrate that the dissymmetry coefficient of a solution of large molecules will, in general, increase with diminishing concentration to a limiting value at infinite dilution. This limiting value, the intrinsic dissymmetry coefficient, $[q]$, is a molecular constant characterising the extension in space of the dissolved molecule. The value of $[q]$ may be readily obtained by extrapolating in a linear manner a plot of $1/q$ against c.

In accordance with theory, the experiments with fractions RT-H and BZO-4 show that the diminution of $[q]$ with increasing concentration is greater the greater the effective solute-solute repulsion; that is, the greater for better solvents (Fig. 3). This behaviour may be pictured roughly as follows: at infinite dilution the q that is determined would correspond to that of the individual molecule. As the solute concentration is increased in a good solvent medium, such as polystyrene in toluene, solute-solute repulsion will *tend* to arrange the solute molecules equidistant from each other, thus bringing a degree of order into the system (such as in a sodium chloride crystal in the extreme case). This " order " produces phase relations and hence destructive interference resulting in a diminution of dissymmetry. Conversely, if the solvent medium is a very poor solvent the solute molecules, upon increasing concentration, will move about in solution in a nearly independent manner which results in the dissymmetry coefficient being nearly independent of concentration, e.g. polystyrene in the butanone-*iso*propanol mixture. This effect is noted in the tobacco mosaic virus also. When dissolved in buffer the repulsive forces (ionic) are minimised and q is independent of concentration. In pure water, however, repulsive forces are dominant and there is a considerable decrease of q with concentration.

As previously indicated, the value of $[q]$ can be interpreted in terms of a characteristic dimension if the molecule is known to correspond in shape to a rod, sphere or random coil. In the case of fractions RT-H and BZO-4, $[q]$ decreases with diminishing solvent power as shown in Table III. (In absence of osmotic pressure data on ethylene dichloride solutions of polystyrene it is not possible to ascertain precisely the solvent ability of this solvent, but it is very probably intermediate between toluene and butanone. The apparent reversal of this ordering for ethylene dichloride and butanone solutions of BZO-4 is considered to be probable experimental error.) This means that the molecular size or extension in space of polystyrene molecules is less in poorer solvents. This is in qualitative agreement with some interpretations of viscosity data.[8] However, a more quantitative comparison is interesting. If polystyrene is assumed to be a randomly coiled, chain-like molecule the value of R can be determined from $[q]$ by use of Fig. 1. The results shown in Table III indicate only a 20 % difference in R between a very good and a very poor solvent. Yet (see Table V) the value of $[\eta]$ in the case of RT-H differs by nearly 500% in the same series of solvents. The variation of $[q]$ demonstrates that the configurations of flexible polymer molecules can be slightly altered by changing the solvent power of the solvent medium. But the large variation of $[\eta]$ with solvent must mean that factors other than molecular size are affecting this quantity. Note that $[\eta]$ varies much more in the series of solvents than it does in the same solvent when the molecular weight is changed six-fold. It may be argued that this is due to temporary, relative solvation of the polymer chain. Thus some toluene molecules may cling for short periods of time to a segment of a polystyrene molecule as it moves through the solution or vice versa. In this manner the size of the polymer segment is effectively increased and the viscosity is greatly enhanced. This effect would be less for poorer solvents and may account for the differences observed. However, since there can be so little real solvation of polystyrene by even the best solvent, toluene, it seems more likely that the predominant factor giving rise to these enormous differences in viscosity is to be found in the nature of the temporary inter- and intra-molecular association of polymer molecules. As the solvent ability of a medium is continuously reduced the heat of dilution will increase; this means that the polymer-polymer contacts become increasingly longer in duration. If the contacts of longer duration occurred only at the ends of the molecule, the effect would correspond to an effectively larger *homologous* molecule and the viscosity would increase. But, of course, the contacts are not at the ends but must occur approximately at random

TABLE IV.—THEORETICAL RATIO OF TRANSVERSE TO LONGITUDINAL SCATTERING FOR POLYSTYRENE FRACTIONS.

Fraction.	Solvent.	Concentration, g./100 ml.					
		1·000.	0·500.	0.250.	0·125.	0·0625.	0·0313.
RT-H	Toluene	0·715	0·695	0·680	0·552	0·470	0·380
,,	Ethylene dichloride	0·855	0·830	0·720	0·565	0·493	0·408
,,	Butanone	0·743	0·726	0·610	0·507	0·463	0·436
,,	Butanone-*iso*-propanol	0·607	0·507	0·468	0·457	0·465	0·463
BZO-4	Ethylene dichloride	—	0·915	0·875	0·850	0·780	—
,,	Butanone	—	0·890	0·820	0·785	0·745	0·0695
,,	Butanone-*iso*-propanol	—	0·823	0·820	0·808	0·767	—
JNZ-7	Butanone	—	0·855	0·795	0·737	0·687	0·665
JNZ-4	,,	—	—	0·677	0·580	0·497	0·463
JNZ-1	,,	—	—	0·630	0·495	0·397	0·350

---

[8] Alfrey, Bartovics and Mark, *J. Amer. Chem. Soc.*, 1942, 64. 1557.

over the entire chain. This situation tends to bring portions of molecules together in lateral arrangement with a consequent clustering and diminution in viscosity. In fact it has been shown[9] that in polyvinyl chloride solutions where this association becomes permanent, the actual molecular weight can increase nearly 100-fold without a significant change in viscosity. Regardless of the explanation the experiments described indicate that $[q]$ is a more reliable measure of relative molecular size than $[\eta]$, inasmuch as $[q]$ is not affected by solvation and other extraneous factors that contribute to the value of $[\eta]$. Moreover $[q]$ may be interpreted in terms of a random coil which is a much better representation of the actual molecule than the often-used axial ratio of a hydrodynamically equivalent ellipsoid.

TABLE V.—MOLECULAR DATA ON POLYSTYRENE FRACTIONS.

Fraction.	Solvent.	Weight Average Molecular Weight M.	Extended Length (A.) L.	Root Mean Square Length R.	Intrinsic Viscosity.	Ratio M/R^2.
RT-H	Toluene	4,000,000	96,000	2370	7·6	—
„	Ethylene dichloride	4,000,000	96,000	2390	5·3	—
„	Butanone	4,000,000	96,000	2100	2·3	0·92
„	Butanone-*iso*propanol	4,000,000	96,000	1900	1·6	—
BZO-4	Ethylene dichloride	1,750,000	42,000	1100	2·16	—
„	Butanone	2,100,000	42,000	1230	1·45	1·15
„	Butanone-*iso*propanol	2,100,000	42,000	990	0·84	—
JNZ-7	Butanone	2,100,000	51,000	1350	1·82	1·14
JNZ-4	„	4,800,000	120,000	2100	3·30	1·08
JNZ-1	„	9,800,000	242,000	2770	4·10	1·27

The determination of $[q]$ for the fractions JNZ-1, 4 and 7 may be used to show that polystyrene molecules in dilute solution are approximately randomly coiled. For an assembly of molecules behaving according to equation (1), R^2 should be proportional to the molecular weight. The testing of this relationship is shown in Table V where it is seen that the ratio M/R^2 is approximately constant for these fractions, and for the two other fractions also if weight average quantities are used. The value of the completely elongated molecule is shown for comparison.

It is of interest to compare the values of R calculated from q with the values obtained from the statistics of a random walk on a diamond lattice.[10] Since we assume the polystyrene molecules to be unbranched, R as calculated from $[q]$ must be considered only as a lower limit. The mean square of the separation of the beginning and end of a random walk in three dimensions is given by

$$R^2 = nl^2 (1 - \cos \theta)/(1 + \cos \theta)$$

where n is the number of steps (valence bonds), l is the length of the step (bond distance) and θ is the angle between successive steps (valence angle). Substituting for n its equivalent M/m, where M is the molecular weight and m is the average weight per valence bond (52 for polystyrene) and 109·5° for θ we obtain 0·092 for R^2/M. This is to be compared with an average value of about 1·1 from Table V. Taking the square roots of the ratios of these two equalities there is obtained

$$R_{\text{expt.}}/R_{\text{theor.}} = 3{\cdot}5.$$

That is, the root mean square of the separation of the ends of the polystyrene molecule is *at least* 3·5 times as large as that which would result from a completely random configuration. Two factors probably account for this effect: one is the finite size of the polystyrene chain which prohibits its crossing back on itself as is possible in a walk; the other is the steric hindrance introduced by the bulky phenyl side-groups which doubtless excludes a significant sector of the 360° that should be available for the orientation of each valence bond normal to the plane formed by the two preceding bonds.

---

[9] Doty, Wagner and Singer, *J. Physic. Chem.* (in press).

[10] See, e.g. W. Kuhn, *Kolloid Z.*, 1934, 68, 2.

It is for these reasons that it is not possible to test directly equation (3) relating the dissymmetry of the root mean square separation of the ends of the coil. However, equation (2) which was derived on the same premises, can be subjected to direct proof by use of tobacco mosaic virus as a model rod-shaped molecule. Data obtained recently, show that $[q]$ for this virus in either pure water or buffer is 0·96. Using the plot in Fig. 1 for rod-shaped molecules in water a $[q]$ of 0·96 is seen to correspond to a L/λ value of 0·67. The wavelength, λ, in solution is 4090 A. The length of the tobacco mosaic virus molecule is then found to be 2750 A. The average value obtained from electron micrographs of this same sample was 2800 A. This would seem to establish the dissymmetry coefficient as a useful molecular constant which can be interpreted accurately in terms of the dimensions of the scattering molecule.

Summary.

The ratio of the intensities of monochromatic light scattered at approximately 45° and 135° from the incident beam traversing polystyrene solutions has been studied in detail. In particular the variation of the dissymmetry coefficient, which is one less than this ratio, has been determined as a function of concentration, solvent ability and molecular weight of the polystyrene fractions. The variation with concentration is shown to depend on solvent ability of the solvent in a predictable manner. The dissymmetry coefficient at infinite dilution, the intrinsic dissymmetry coefficient, can be accurately evaluated and is recognised as a molecular constant. This intrinsic dissymmetry coefficient, $[q]$, is most useful as a characterisation of particle size for it is a weighted measure of the extension in space of the material composing the particle or molecule. A characteristic dimension can be obtained from $[q]$ if the scattering particle approximates in form a sphere, rod or random coil. By this means the root mean square distance between the ends of the polystyrene molecule are calculated for fractions whose molecular weights were determined by light scattering (and confirmed by osmotic pressure measurements). These root mean square distances were found to be proportional to the square root of the molecular weight, thus showing that the coiling of the molecules follows a random statistics, at least in butanone solutions. However, the actual root mean square distances were at least 3·5 times greater than those calculated for an ideal case; this was considered to be due to the finite thickness of the molecule and the steric hindrance preventing free rotation. In a series of different solvents the intrinsic viscosity of a polystyrene fraction was found to differ by about 500% whereas the root mean square distance varied only about 20%; this is believed to be due to the influence of solvent-solute interaction on details of the molecular configuration. These observations emphasise the utility of the intrinsic dissymmetry coefficient as a measure of particle size in comparison with intrinsic viscosity.

These data on tobacco mosaic virus were obtained in collaboration with Dr. G. Oster of the Rockefeller Institute for Medical Research. It is planned to publish a complete account of the work elsewhere.

Polytechnic Institute of Brooklyn,
Brooklyn 2, N.Y., U.S.A.

GENERAL DISCUSSION *

Prof. P. Putzeys (*Louvain*) said: Molecular weight determinations on protein molecules in aqueous solution have been made since 1935 by the light scattering method.[1, 2] The largest molecules so far examined are those of the hæmocyanin of *Helix Pomatia* with a molecular weight of nearly 8,000,000 and a length of 890 A. This length appears to be just below the limit of 1/20 of the wavelength and, judging from depolarisation measurements,[4] there does not appear a convincing dissymmetry in the intensity of the light scattered in the forward and backward directions.

From the information available to me it is not quite clear how far light absorption, intermolecular interference and depolarisation intervene in modifying the dissymmetry of the scattered light observed by Doty, Affens and Zimm. The depolarisation increases rapidly for angles of observation different from 90° and its contribution to the total scattering becomes large, and this phenomenon is furthermore complicated by secondary scattering.[3, 4]

Dr. G. Gee (*Welwyn Garden City*) said: Doty's estimate of the root mean square length of polystyrene molecules in solution affords the first direct experimental test of the effect of solvent on the molecular configuration of a dissolved polymer. My late colleague, W. J. C. Orr, carried out some approximate calculations of the number of polymer liquid contacts to be expected in an infinitely dilute solution and the way in which this number depends on the heat of mixing. An account of this work has been accepted by the Faraday Society for publication, and I need only record here his conclusions, which were: (1) in an athermal solution the number of polymer liquid contacts would be roughly 60 % of the number corresponding with the completely outstretched chain; (2) this number would be only slightly diminished by any endothermic mixing consistent with the liquid being a solvent. The first point is to be compared with our experimental estimate of 77 % for rubber in benzene;[5] the second is evidently consistent with Doty's observations.

Dr. Magat (*Paris*) said: The small alteration in the size of a swollen polymer molecule in changing from a good to a poor solvent observed by Doty and his co-workers probably means the end of the assumption that the viscosity variation is due to a volume increase by "hydration." Hence I would welcome comments on a suggestion put forward by Carter, Scott and myself. We suggested that the actual viscosity observed arises from two competitive mechanisms of flow. In the first mechanism, the polymer molecule may be considered as a dense sphere moving together with the entrapped liquid, in which case the Einstein formula applies, and $[\eta] \sim M^{\frac{1}{2}}$. In the second mechanism, the increase in the size of the molecule is enough to allow some lines of flow to go through the molecule and to act on each individual link, as in the case considered by Huggins. For this mechanism $[\eta] \sim M$. The actual molecular weight dependence is intermediate, being about $[\eta] \sim M^{\frac{3}{4}}$ in a good solvent and $[\eta] \approx M^{\frac{1}{2}}$ in a poor one. This change may be explained by the co-existence of the two mechanisms and would enable us to understand how a relatively small change in the swelling of the dissolved molecule (making the second mechanism possible, or increasing its contribution) induces a large change in the viscosity.

Dr. P. M. Doty (*Cambridge*) said: Measurements such as those of Alfrey, Justice and Nelson or our measurements of the viscosity of polystyrene in different solvents are consistent with the hypothesis of the extension of polymer molecules being quite sensitive to relatively small changes in heats of dilution. However, the changes in heats of dilution for polystyrene solutions are particularly small for measurements, which we have recently completed on the temperature dependence of the osmotic pressure of polystyrene solutions, show that over the whole range of solvents for polystyrene, α ($= \Delta\overline{H}_1/v_2^2$) varies less than 50 cal. This is only about one-fifth of the range found in polyvinyl chloride solutions. Consequently, for polystyrene solutions it would appear that the entropy changes coincident with the large alterations of molecular size postulated are incompatible with the small range of heat changes available.

---

* On preceding paper.

[1] Putzeys and Brosteaux, *Trans. Faraday Soc.*, 1935, **31**, 1314.

[2] Putzeys and Brosteaux, *Med. Kon. VI Academie Wetensch.*, 1941, **3**, no. 5.

[3] Lontie, *ibid.*, 1944, **6**, no. 11.

[4] Lotmar, *Helv. Chim. Acta*, 1938, **21**, 953.

[5] Gee and Orr, *Trans. Faraday Soc.*, 1946, **42**, 507.

The alternative explanation offered at the end of our paper is equally adequate to explain the observed results. This may be a clear case of our knowledge still being in the " non-theory sensitive " stage, to use Prof. Bernal's expression. However, this later explanation is supported by the light scattering experiments and is consistent with the theoretical considerations of the late Dr. Orr. Moreover, we have found further support for this general idea of intrinsic viscosity behaviour from recent work on the stable association of polyvinyl chloride molecules in solution. Here it was found that stable association could take place to the extent of increasing the weight average molecular weight one-hundred fold, yet the intrinsic viscosity was only minutely changed. Dissymmetry measurements showed that the molecular size was little changed. This is an example of how the detailed nature of the intermolecular contacts—in this case stable instead of temporary—and not necessarily the size or mass of the polymer molecule determines the intrinsic viscosity.

Dr. F. Eirich (*Cambridge*) said : Concerning Dr. Magat's questions, the problem of the relation between particle dimensions, solvent immobilisation and the exponent in the viscosity law have been rather fully discussed by Kuhn [6] and also by R. Simha.[7] Model experiments have been carried out by myself in collaboration with Huber and Mark on spheres, rods and plates,[8] showing that large amounts of liquid do in fact move with hollow or porous particles whenever the pressure differences across the particle are too small to cause appreciable draining. This proved the existence of the frequently postulated " immobilisation," or hydrodynamic solvation, which of course expresses itself mechanically only and is non-existent in the thermodynamical sense.

As for Dr. Doty's problem of the discrepancy between the relatively small change in molecular asymmetry as revealed by light scattering, and the large changes in intrinsic viscosity, on varying the solvent, may I refer first to an earlier discussion of this topic during the Faraday Meeting.[9] I agree with Dr. Doty's remarks that changes in molecular coiling may well affect viscosity in more ways than just by altering the geometry. On loosening up in a better solvent there are first the opposed effects of decreasing solvent inclusion and increased hydrodynamic molecular interaction, i.e., a higher concentration effect. Incidentally, this antagonism together with certain types of association, may lead to compensation and may thus explain the opposite effect also observed—that of a very small dependence of the viscosity on solvent composition.

Further, there is the possibility of more numerous direct molecular contacts already referred to. In this respect I wish to emphasise the great service the light scattering method may render in future by its linear dependence on the concentration in a range in which the viscosity may rise more steeply, and in particular by its power of revealing whether there are really single molecules in a given solution or not. The same possibilities have also been opened by measuring the temperature coefficient

[6] Kuhn, *Helv. Chim. Acta*, 1943.
[7] Simha, *J. Chem. Physics*, 1945.
[8] Eirich, Huber and Mark, *Papierfabrikant*, 1937.
[9] *Trans. Faraday Soc.*, 1944, **40**, 277.

of osmotic pressure, as now introduced by Dr. Mark's school with striking results.

Thirdly, I want to underline the fact that with the opening of molecular coils the hydrodynamically effective dimensions may change considerably without large changes in the overall asymmetry. This is provided for implicitly by Simha's, and explicitly by Kuhn's, latest treatment of this question, quoted earlier above. According to them :

$$\eta_{sp}/c \sim (L/D),\ L/D \sim M^a;\ \text{or}\ \ \eta_{sp}/c \sim \overline{h}^2 L/Z$$

respectively; $h(=R)$ is the root mean square end-separation, L the effective hydrodynamic length, Z the degree of polymerisation, a = constant. It will be seen that changes in a of a magnitude which may easily occur, e.g., from 0·5 to 0·8, will affect the viscosity greatly by changing the effective axial ratio without, however, coinciding with the geometrical one except in case of a compact ellipsoid. The duality is even better shown by considering Kuhn's formula where L is formally independent of the mean square end-separation and is subject to much greater alterations than the latter. Herein, I think, Dr. Doty will find the explanation for his results which in turn are very valuable, in that they demonstrate that large differences in intrinsic viscosities need not necessarily reflect large changes in molecular asymmetry.

Mr. F. W. Peaker (*Aberdeen*) said : In order to measure the weight mean average molecular weight by light scattering methods, it is necessary to measure the turbidity and refractive index of the solution at varying concentration. These two quantities are the only experimental values required to give a molecular weight figure. No reference to any other instrument, for example, an osmometer or ultracentrifuge is necessary. This is important since the method is the most ready means of obtaining an absolute value of the weight mean molecular weight.

It has been stated by Doty and co-workers,[10] that, although satisfactory agreement was obtained between osmotic molecular weights and the value deduced from light scattering measurements for cellulose acetate solutions, this was not so for polystyrene solutions. We are not told the order of this discrepancy, but it could be attributed to either of the following causes: (*a*) the original treatment due to Debye is insufficiently rigid, or (*b*) that the correction to be applied to the scattered light due to its dissymmetry is in error. The former might be true if the divergence is large, and the latter if small.

This matter has been overcome in the present paper by the use of a polystyrene solution of low molecular weight as the calibration standard for the turbidimeter, and the use of a spectrophotometer to determine the extinction coefficient of this solution.

In view of the importance attached to the dissymmetry coefficient in this work, perhaps Dr. Doty could explain the cause and nature of these earlier faults in the method. This point is of special importance in the establishment of the light scattering method as an absolute method of determining the weight mean average molecular weight.

Dr. P. M. Doty (*Cambridge*) said: The light scattering method is an absolute method of determining weight average molecular weights. The instrument one uses to measure the intensity of the transversely scattered light may be calibrated with a solution whose scattering is symmetrical (Rayleigh) and sufficient to permit an accurate absolute determination of its turbidity by a transmission measurement. Earlier difficulties were found to be due to the magnesium carbonate reflection calibration. We have replaced this method with a transmission measurement as described.

The first determination of molecular weights by light scattering are those of Prof. Putzeys and Dr. Brosteaux on several proteins. In this work they had to evaluate a constant by means of measurements on a protein of well-known molecular weight. This constant can now be calculated from known constants, as Debye has pointed out, thus eliminating the semi-empirical nature of the method.

---

[10] *J. Amer. Chem. Soc.*, 1946, **68**, 159.

MUTUAL SOLUBILITY OF LIQUIDS.

By A. E. van Arkel.

Received 25th July, 1946.

The process of swelling, e.g. of rubber in benzene is nearly related to the solution of a liquid in benzene. In both cases benzene molecules must penetrate into the second phase, and molecules of the latter must go over into the benzene phase.

Before attacking the problem of the kinetics of the swelling process, we should know what changes in the thermodynamical properties are connected with this process.

The change in energy can be calculated in a manner analogous to the calculation of the heat of solution of two liquids. As both the heat of solution M and the heat of vaporisation V are directly connected with the cohesion energy, the heat of solution of two liquids may be expressed as a function of the heat of vaporisation of both liquids.

A formula of this type was given already by Hildebrand:

$$M = -\alpha(1 - \alpha)(\sqrt{V_1} - \sqrt{V_2})^2$$

expressing M in terms of the heat of vaporisation, V, and the mol. fraction α, of both compounds.[1]

We will, however, show that this expression cannot describe all the features of mutual solubility. As a matter of fact, it turns out that Hildebrand's formula holds very well or at least is a good approximation for the heat of solution of perfectly non-polar compounds especially when the molecules of both compounds are not of very different form or size.

The London energy between two particles A_1 and A_2 in contact is:

$$E_{1.2} = -c\frac{\phi_1\phi_2}{\phi_1 + \phi_2}\frac{\alpha_1\alpha_2}{(R_1 + R_2)^6}.$$

When ϕ_1 and ϕ_2, R_1 and R_2 have not greatly different values, we may write:

$$E = -C\sqrt{\phi_1\phi_2}\frac{\alpha_1\alpha_2}{2^6R_1^3R_2^3} = -\sqrt{E_1E_2}$$

in which formula E_1 and E_2 are the energies of either two particles of the first type, or two particles of the second type, in contact.

In 1 mol. of a close-packed liquid, A_1, where each particle is surrounded by p neighbours, the total number of contacts is:

$$\frac{pN}{2} \text{ and the total cohesion energy } L_1 = -\frac{pN}{2}E_1.$$

In a mixture of α mol. A_1 and $(1 - \alpha)$ mol. A_2, the total energy is:

$$L_{1.2} = -\frac{\alpha^2pN}{2}E_1 - \frac{(1 - \alpha)^2pN}{2}E_2 - \frac{2\alpha(1 - \alpha)pN}{2}E_{1.2}$$

and the energy change when the mixture is formed:

$$M = -\alpha\frac{pN}{2}E_1 - (1 - \alpha)\frac{pN}{2}E_2 - L_{1.2}$$

$$= -\alpha(1 - \alpha)\frac{pN}{2}(\sqrt{E_1} - \sqrt{E_2})^2 = -\alpha(1 - \alpha)(\sqrt{L_1} - \sqrt{L_2})^2.$$

---

[1] Hildebrand and Wood, *Solubility* (N.Y, 1936).

The fact that real liquids are not close-packed does not greatly affect the result; in this case R is to be replaced by a mean value $\bar{R}$ which will be the same function of R for A_1, A_2 and the mixture. If the molecules are of different size, the formula becomes

$$M = -\frac{N_1N_2}{\beta(N_1+N_2)} \cdot (\sqrt{L_1} - \sqrt{\beta L_2})^2$$

where β is the quotient of the surfaces of the two kinds of molecules.[2] The formula is invalid, however, when in one of the compounds the cohesion energy is mainly due to another type of force, *e.g.* to dipole interaction.

Let us assume for a moment that just as for the London energy in the case of dipole-dipole interaction, giving rise to energies K_1 and K_2, the cohesion energy between different particles is equal to $\sqrt{K_1K_2}$.

In that case the heat of mixing will be equal to

$$M = -\alpha(1-\alpha)\{(\sqrt{L_1} - \sqrt{L_2})^2 + (\sqrt{K_1} - \sqrt{K_2})^2\},$$

which in the extreme case of water, which has practically only dipole-energy ($L_1 = 0$) and heptane with only London energy ($K_2 = 0$) becomes

$$M = -\alpha(1-\alpha)(L_2 + K_1);$$

the heat of mixing then attains a large negative value.

This formula expresses the well-known fact that non-polar-liquids do not mix with polar ones, and that the solubility increases with decreasing difference between the two. It gives, however, no explanation of the occurrences of positive heat of mixing of two polar compounds, as e.g. in the system $C_2H_5OH—H_2O$. In some cases these positive heats of mixing will be due to the formation of chemical compounds as e.g. in the systems $H_2O—NH_3$, $H_2O—HCl$, etc., but formation of compound-molecules is not so evident, for the system $CHCl_3$-acetone that has a positive heat of mixing as well.

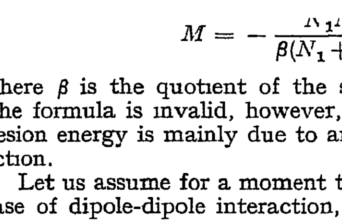


FIG. 1.

In the derivation of the formula it was assumed that although there is a change in the energy, the distribution of the molecules in the mixture is still statistical. This is no longer true, when in the molecules the dipoles are so situated that in one the positive pole, in the other the negative, is nearer to the surface.

In that case configurations of a special type (Fig. 1) will have a dipole-interaction energy $F \gg \sqrt{K_1K_2}$ and, moreover, such configurations will occur much more frequently than we should expect in the case of a purely statistical distribution. Energy and entropy now become very complicated functions of the concentration and the properties of both compounds.

In this case closed regions of two-phase equilibria in systems of two liquids may be formed (H_2O-alkyl-pyridines, H_2O-nicotine).

It is interesting that in systems with a closed two-phase region this region always lies in the concentration region 100 % mol. H_2O-50 % mol. alkyl-pyridine. Clearly the pairs of pyridine-water molecules are so stable, that they may be described as molecules of the hydrate. Now the London energy of the hydrate is of the same order of magnitude as that of pyridine, and the same holds for the dipole-interaction energy. Water and the hydrate, however, have a widely different London and dipole energies. Thus water is not miscible in all proportions with the hydrates, the latter, however, are miscible in all proportions with the alkyl-pyridines.

---

[2] Staverman, *Rec. trav. Chim.*, 1937, 56, 885.

The phase diagram indeed is of the type of Fig. 2. A theory of systems with positive heat of mixing is given by Staverman.[3]

To discuss the mutual solubilities of liquids we must know what part of the cohesion energy is due to London-v.d. Waals forces, and what part is caused by dipole-dipole interaction. We have further to consider the Debye energy (dipole-induction effect) which, however, in general, is so small that it may be neglected.

An estimate of the dipole energy can be made in the following way: compounds CH_4, CF_4, CCl_4 with $\mu = 0$ will have London energy only. If intermediate compounds CH_3F, CH_2Cl_2, etc., would have no dipole energy, their heat of vaporisation could be calculated with the aid of the following relation:

$$V_{CH_3F} = \tfrac{1}{4}V_{CF_4} + \tfrac{3}{4}V_{CH_4}.$$

Which, as all these compounds have nearly the same Trouton constant, may be written in the form:[4]

$$T_{CH_3F} = \tfrac{1}{4}T_{CF_4} + \tfrac{3}{4}T_{CH_4},$$

where T is the boiling-point[4] in °K.

For all halides CH_3X the boiling-point calculated in this way is ~ 80° lower than the experimental. This means that the energy due to dipole interaction for these compounds is 1·7 kcal.; for CH_2X_2 we find ~ 1·8 kcal., for CHX_3 ~ 1 kcal.

In the series (CH_4, CH_3Cl, CH_2Cl_2, $CHCl_3$, CCl_4), CH_3Cl will show the greatest miscibility with water: it must be more soluble than (hypothetical liquid) CH_4 because of its dipole energy: with increasing Cl content the dipole energy decreases, the London energy increases, so the solubility both of water in the halides and vice versa must decrease.

Solubility in g./l.:

CH_4	CH_3Cl	CH_2Cl_2	$CHCl_3$	CCl_4
0	4·1	—	0·82	0·08

In the sequence CH_3F, CH_3Cl, CH_3Br, CH_3I the solubility decreases, due to the increase of the London energy. The solubility data are too uncertain for comparison with the calculated values: for the moment we can only expect to find qualitative agreement between theory and experiment.

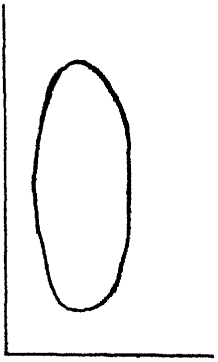


FIG. 2.

It is to be emphasised that there is no direct simple relation between dipole energy and total dipole moment.

If in a large molecule there are two dipole groups each of them will give an independent contribution to the cohesion energy. We find, indeed, that the difference in boiling point between a normal hydrocarbon and its 1-chloro derivative is the same as that between the latter and the *n*-dichloro compounds, and this will remain true even when the dipoles have fixed positions in space with opposite direction as for instance in the *p*-benzene halides.

If the two Cl atoms come near together, as in $ClCH_2 . CH_2Cl$ there will be interaction between the two, and when both Cl atoms are bound to the same C atom, they will behave as one single group with dipole $\vec{\mu} = \Sigma\vec{\mu}_1$. So when a number of dipole groups are at a great distance from each other the contribution to the cohesion energy will be proportional to $\Sigma \mu_i^2$ whereas this contribution will be proportional to $(\Sigma \vec{\mu_i})^2$ when the groups are bound to the same C atom.

---

[3] Staverman, *Physica*, 1937, 4, 1141; *Rec. trav. Chim.*, 1941, 60, 827.

[4] Van Arkel, *Trans. Faraday Soc.*, 1934, 30, 698.

Changes in boiling points which can be used as a measure for the cohesion energies go parallel to solubilities, as was shown for the chloroethanes.[5]

With increase of the Cl content the solubility in water decreases as a result of the increasing polarisability, but in each pair of isomers the solubility is greater and the boiling point higher for the compound that has two separate dipole groups in the molecule. It is the dipole energy, rather important in comparison with the total London energy, that makes $CH_2Cl . CH_2Cl$ the useful solvent it really is.

It is a well-known fact, that aromatic compounds are more soluble in water and polar liquids than aliphatic compounds. This difference cannot, according to the formula for the cohesion energy, be attributed to the greater polarisability of the aromatic nucleus, as this effect should decrease the solubility of the aromatic compounds in water.

The effect must be due to the dipole moments of the C—H bonds, which, although probably only of the order of $0 \cdot 4 \times 10^{-18}$, make a contribution to the cohesion energy which is nearly equal to that of the dipole of a C—F or C—Cl bond. (For all fluoro compounds with dipole moment $\mu = 0$ (CF_4, C_2F_6, etc.) the boiling point is nearly equal to that of the corresponding hydrocarbons, from which it follows that the London energy is the same for H and F.)

The difference between CH_4 and CH_3F, therefore, is due only to the difference in dipole energy ($\Delta T = 80°$). The boiling points of C_6H_5F and $C_6H_4F_2$ are nearly equal to that of C_6H_6; in other words, these compounds have the same dipole energy as well. In the substitution H → F in benzene the —C—F dipole replaces the equally active dipole —C—H, that is to say a dipole that gives nearly the same contribution to the cohesion energy.

The polar character of aromatic compounds is shown by the solubility of the compounds in liquid SO_2, which has just the right proportion of dipole cohesion energy to dissolve aromatic compounds, but does not mix with non-polar aliphatic hydrocarbons. Unsaturated hydrocarbons show the same effect: the groups C═C and C≡C behave as polar groups, both in their effect on boiling-points and solubility, but at present it is not possible to explain in detail in what manner these effects are influenced by substitutions of the H atom in unsaturated compounds by other atoms or groups.

Anorganisch Chemisch Laboratorium
der Rijks-Universiteit,
Hugo de Grootstraat 27, Leiden, Holland.

---

[5] Staverman, *Rec. Trav. Chim.* 1941, **60**, 836; Van Arkel and Vles, *ibid.*, 1936, **55**, 407.

ABSORPTION OF WATER BY POLYMERS: ANALYSIS IN TERMS OF A SIMPLE MODEL.

By A. J. Hailwood and S. Horrobin.

Received 15th *August*, 1946.

The absorption of water by the older textile fibres—wool, silk and cotton—has long been known to be characterised by absorption isotherms of sigmoid shape. Other natural fibres, and all the newer synthetic fibres for which absorption isotherms have so far been determined, show the same characteristic, as do many non-fibrous proteins.

Obviously neither the empirical isotherm of Freundlich nor the Langmuir isotherm for monomolecular adsorption can fit such cases, and the more recent isotherm of Brunauer, Emmett and Teller[1] which extends Langmuir's treatment to cover multimolecular adsorption has also only partial success.[2]

No attempt will be made here to review the extensive literature on adsorption in general, or on the absorption of water by textiles, or by proteins, in particular. For textiles, there is a recent review by Carlene.[3] In the case of proteins, fibrous and otherwise, attempts have been made to account for the shape of the experimental isotherms by postulating two kinds of combined water, attached respectively to the polar side chains and to the polar groups in the polypeptide chains, as well as a final condensation of liquid water in the pores of the fibres.

In this paper, a different model is chosen. The adsorbed water is assumed to exist in two states, first water in simple solution, and secondly water combined to form a hydrate with a definite unit of the fibre molecule. It is then simply assumed that the three species present in the solid phase—dissolved water, unhydrated molecules, and hydrated molecules—form an ideal solid solution.

This assumption of ideality in the solid phase allows the absorption isotherm and also the equation connecting volume and amount of adsorbed water to be explicitly derived, and tested by comparison with experimental data.

In all those cases in which both absorption and desorption isotherms have been experimentally determined, with the exception of Nylon, there is a marked hysteresis which appears to be real. The simple model adopted does not account for the hysteresis. A possible modification is suggested, but has not yet been developed, and there are as yet no data of sufficient precision to warrant the application of an equation with the additional constants which would be required by such an extension.

Derivation of the Isotherm.

The model chosen postulates that the absorption of water vapour by polymeric substances is governed by two sets of equilibria (1) the formation of a solid solution of water in the polymer, and (2) the formation of hydrates between water and definite units of the polymer molecule. It is further assumed that the mixture of polymer, polymer hydrates, and dissolved water forms a single solid phase. The dissolved water clearly has no tendency to separate from this mixture. There might, however, be a tendency for the polymer and the polymer hydrates to separate into two or more phases, in the sense that this change would be accompanied by a decrease in free energy, but in fact the immobility and entanglement of the molecules would cause such a separation to be so slow that the system would in practice be quite indistinguishable from a single phase.

The two component system, water-polymer, will exist then in two phases, the solid solution and water vapour, and will therefore be bivariant. The curve illustrating the relation between water content of the polymer, and the external water vapour pressure will thus be smooth and free from discontinuities. The experimental curves are in fact of this sort.

The dissolved water molecules in the solid phase are assumed to be mobile. This is sufficient to secure equilibrium between (*a*) the dissolved water and the external water vapour and (*b*) the dissolved water and the anhydrous and hydrated polymer molecules, in spite of the substantial immobility of these molecules themselves.

At constant temperature and total pressure, and at a particular pressure of the external water vapour, let a' be the activity of the dissolved water in the solid phase and w_i the activity of the i^{th} hydrate of the polymer, then

---

[1] Brunauer, Emmett and Teller, *J. Amer. Chem. Soc.*, 1938, **60**, 309.

[2] Bull, *ibid.*, 1944, **66**, 1499.

[3] Carlene, *J. Soc. Dyers Col.*, 1944, **60**, 232.

$$\frac{w_i}{w_{i-1}a'} = K_i\,; \qquad i = 1, 2, \ldots n \quad . \quad . \quad . \quad (1)$$

where n represents the highest hydrate formed, and K_i are the appropriate equilibrium constants, and, from (1)

$$w_i = w_0\,(a')^i\,K_1K_2 \ldots K_i \quad . \quad . \quad . \quad . \quad (2)$$

where w_0 is the activity of the unhydrated polymer.

Now if a is the activity of the water in the vapour phase, and K is a constant, we have

$$a' = Ka \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

so

$$w_i = w_0(Ka)^iK_1K_2 \ldots K_i \quad . \quad . \quad . \quad . \quad (4)$$

On our assumption that the solid solution is ideal, the activities of the species present may be equated to their mole fractions. In the vapour phase, it will be sufficient, at the low vapour pressures involved, to equate the activity to the fractional humidity. Thus if h is the relative humidity, expressed, as is customary, as a percentage

$$a = h/100. \quad . \quad . \quad . \quad . \quad . \quad (5)$$

If x_i is the number of moles of the i^{th} hydrate, and s is the number of moles of free dissolved water in the solid phase

$$w_i = x_i\Big/\Big(s + \sum_{0}^{n} x_i\Big) \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

$$a' = s\Big/\Big(s + \sum x_i\Big) = Ka. \quad . \quad . \quad . \quad . \quad (7)$$

Now the total water, W, in the solid phase is $s + \sum_{i=1}^{n} ix_i$, and from (4) and (6)

$$x_i = x_0(Ka)^iK_1K_2 \ldots K_i \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$\therefore \qquad W = s + x_0\sum_{i=1}^{n} i\,(Ka)^iK_1K_2 \ldots K_i \quad . \quad . \quad . \quad . \quad (9)$$

or, since the total number of moles of polymer is given by

$$\sum_{i=0}^{n} x_i = x_0 + x_0\sum_{i=1}^{n} (Ka)^iK_1K_2 \ldots K_i \quad . \quad . \quad (10)$$

the concentration, c, of water molecules per mole of polymer is

$$\frac{s}{\sum\limits_{i=0}^{n} x_i} + \frac{\sum\limits_{i=1}^{n} i\,(Ka)^iK_1K_2 \ldots K_i}{1 + \sum\limits_{i=1}^{n} (Ka)^iK_1K_2 \ldots K_i} \quad . \quad . \quad (11)$$

But from (7)

$$s = Ka\sum_{i=0}^{n} x_i/(1 - Ka) \quad . \quad . \quad . \quad . \quad (12)$$

so that

$$c = \frac{Ka}{1 - Ka} + \frac{KaK_1 + 2(Ka)^2K_1K_2 + \ldots .}{1 + KaK_1 + K_1K_2(Ka)^2 + \ldots} \quad . \quad . \quad (13)$$

Let $\alpha h = Ka$, then from (5) $100\alpha = K$, and let $\beta_i = K_1K_2 \ldots K_i$. Also let M be the "molecular weight" of the operative polymer unit.

Then the isotherm may be written in a form permitting comparison with the experimental data as usually determined:

$$\frac{Mr}{1800} = \frac{\alpha h}{1 - \alpha h} + \frac{\beta_1 \alpha h + 2\beta_2 \alpha^2 h^2 + \ldots}{1 + \beta_1 \alpha h + \beta_2 \alpha^2 h^2 + \ldots} \quad . \quad . \quad (14)$$

where r is the regain—that is the total grams of water absorbed per 100 grams of dry polymer. It will be seen that the first fraction on the right-hand side represents the fraction of free dissolved water, and the various terms of the second fraction represent the fractions of water present as monohydrate, dihydrate, etc.

In fitting the isotherm to the available data, it has been found sufficient to postulate only a monohydrate, when the isotherm reduces to

$$\frac{Mr}{1800} = \frac{\alpha h}{1 - \alpha h} + \frac{\alpha \beta h}{1 + \alpha \beta h} \quad . \quad . \quad . \quad (15)$$

Comparison with Experimental Data.

The fitting of the isotherm of equation (15) to experimental data involves the calculation of the " best " values of the three constants α, β and M. To facilitate computation, the equation was transformed to the form

$$A + Bh - Ch^2 = h/r \quad . \quad . \quad . \quad . \quad (16)$$

and the best values of A, B and C were determined by the method of least squares; α, β and M were then found from their known algebraic relations to A, B and C.

The isotherm has been fitted in this way to a number of sets of data on wool from different sources,[2, 4] and to data for hair,[5] silk,[6] Nylon [3] and cotton [7] as well as to all the non-fibrous proteins for which data are given in the comprehensive paper by Bull.[2]

It will be illustrated here by two sets of data for wool, and one set each for hair, silk, Nylon, and cotton. These results are typical of the fits obtained in all the cases so far examined. They are shown in Tables I to III. It will be seen that the fit is quite good.

From the values of α and β, the standard free energy changes governing the solution of water in the polymer, and the combination of dissolved water with unhydrated polymer to form the hydrate, can readily be calculated. Thus

$$\Delta G^\circ_{sol} = -RT \log_e 100\alpha \quad . \quad . \quad . \quad (17)$$

is the free energy change corresponding to the transfer at temperature T° K. of 1 mole of water from the external, ordinary liquid form, or from the saturated vapour, to that hypothetical state implied in equating the activity a' in the solid solution to the mole fraction of water present. It is therefore the state in which the water, in the solid solution, has a mole fraction of unity. In this state the vapour pressure of the water is greater than that of pure liquid water in the ratio 1 : 100α.

Further

$$\Delta G^\circ_{chem} = -RT \log_e \beta \quad . \quad . \quad . \quad (18)$$

is the change in free energy when one mole of dissolved water in its standard state in solution combines with the polymer to form one mole of monohydrate—also in its standard state when its mole fraction is unity; and finally

$$\Delta G^\circ_{total} = \Delta G^\circ_{sol} + \Delta G^\circ_{chem} = -RT \log_e 100\alpha\beta \quad . \quad . \quad (19)$$

is the change in free energy when one mole of ordinary liquid water, or of saturated vapour, reacts with dry polymer to form one mole of monohydrate.

---

[4] Speakman, *J. Soc. Chem. Ind.*, 1930, **49**, T.305.
[5] Chamberlain and Speakman, *Z. Elektrochem.*, 1931, **37**, 374.
[6] Goodings and Turl, *J. Text. Inst.*, 1940, **31**, T.69.
[7] Urquhart and Williams, *ibid.*, 1924, **15**, T.433.

The values of M, ΔG°_{sol}, ΔG°_{chem} and ΔG°_{total} are given in Table III for the materials of Tables I-II and for a number of others also.

TABLE I.

Relative Humidity (%).	Regain (%).		Relative Humidity (%).	Regain (%).	
	Observed.	Calculated.		Observed.	Calculated.
		Wool at 25° C. (Speakman [8])			
10	4·08	4·03	64·5	14·83	14·58
25	6·96	7·13	74·5	17·37	17·56
34·7	8·68	8·73	89·7	23·73	24·57
49·5	11·50	11·30	100·0	33·90	32·98
		Human Hair at 22·2° C. (Chamberlain and Speakman)			
8·5	3·93	3·85	63·0	14·84	14·73
24·2	7·32	7·44	85·8	22·55	22·34
40·4	10·20	10·22	100·0	31·18	31·44
		Silk at 25° C. (Goodings and Turl)			
3·5	1·1	1·2	74·5	11·6	11·8
9·9	2·4	2·5	79·9	12·5	13·5
15·1	3·4	3·3	89·8	17·2	18·2
19·5	3·8	3·8	97·5	25·4	24·7
34·5	5·7	5·4	99·4	30·5	27·1
64·5	9·5	9·4			
		Cotton at 25° C. (Urquhart and Williams)			
0·5	0·37	0·24	68·9	7·31	7·35
1·1	0·59	0·51	74·2	8·21	8·27
1·8	0·87	0·76	79·6	9·20	9·43
4·6	1·37	1·54	84·2	10·14	10·70
8·6	1·84	2·11	87·7	11·56	11·91
13·6	2·32	2·62	92·0	13·27	13·80
27·9	3·48	3·62	92·9	13·94	14·28
37·0	4·36	4·21	96·0	15·60	16·19
46·4	5·06	4·90	97·7	18·04	17·64
50·5	5·38	5·24	100·0	22·97	19·54
62·1	6·52	6·43			

TABLE II.

Relative Humidity (%).	Wool (Bull). Regain (%).				Stretched Nylon (Bull). Regain (%).			
	25° C.		40° C.		25° C.		40° C.	
	Observed.	Calculated.	Observed.	Calculated.	Observed.	Calculated.	Observed.	Calculated.
5	2·97	2·64	2·50	2·05	0·58	0·50	0·49	0·43
10	4·25	4·29	3·88	3·60	0·83	0·88	0·68	0·78
20	6·25	6·46	5·65	5·86	1·36	1·49	1·16	1·35
30	8·22	8·14	7·50	7·76	1·92	1·99	1·71	1·85
40	9·90	9·72	9·36	9·39	2·48	2·48	2·24	2·34
50	11·43	11·43	10·90	11·15	3·10	3·00	2·80	2·86
60	13·47	13·42	12·66	13·10	3·72	3·60	3·36	3·45
70	15·61	15·91	14·70	15·37	4·48	4·33	4·06	4·16
80	18·05	19·21	17·16	18·26	5·28	5·29	4·88	5·09
90	22·54	23·91	20·97	22·10	6·57	6·64	6·03	6·38
95	29·15	27·14	25·75	24·40	7·70	7·55	6·65	7·24

Speakman, *Trans. Faraday Soc.*, 1929, 25,

TABLE III.

Substance.	*M.*	Calories.		
		$+\Delta G^{\circ}_{\text{sol.}}$	$-\Delta G^{\circ}_{\text{chem.}}$	$-\Delta G^{\circ}_{\text{total.}}$
	(A) at 25° C.			
Wool [8]	218·9	164	1,307	1,143
Wool [4] 80's Merino	233·9	162	1,444	1,282
Wool [4] Oxford Down	218·2	166	1,325	1,159
Wool [2]	219·2	174	1,400	1,225
Human Hair [5]	202·4	187	1,313	1,126
Silk [6]	383·1	107	1,344	1,237
Silk [2]	395·3	107	1,574	1,467
Nylon [9]	751·6	184	1,484	1,300
Stretched Nylon [2]	755·0	178	1,012	834
Unstretched Nylon [2]	805·0	110	932	822
Cotton [7]	556·3	106	1,736	1,630
	(B) at 40° C.			
Wool [2]	197·1	236	1,197	961
Silk [2]	395·0	138	1,504	1,367
Stretched Nylon [2]	795·6	194	934	739
Unstretched Nylon [2]	696·6	142	611	469

Calculation of Heats and Entropies of Reaction.

In his paper Bull gives absorption data for the temperatures 25° C. and 40° C., which are sufficiently close to allow the heats and entropies of reaction to be computed by the known methods. The results are given in Table IV.

Since the solid solution is ideal, the heat of mixing of its constituents in their standard states for the solid phase is zero. Thus by mixing pure polymer and liquid water in such proportions as to be in equilibrium with vapour at humidity *h*, the change in enthalpy is

TABLE IV

Substance.	Calories.			Entropy Units.		
	$-\Delta H^{\circ}_{\text{sol.}}$	$-\Delta H^{\circ}_{\text{chem.}}$	$-\Delta H^{\circ}_{\text{total.}}$	$-\Delta S^{\circ}_{\text{sol.}}$	$-\Delta S^{\circ}_{\text{chem.}}$	$-\Delta S^{\circ}_{\text{total.}}$
Wool [2]	1,056	5,416	6,473	4·13	13·48[4]	17·61
Silk [2]	498	2,960	3,457	2·03	4·65	6·69
Stretched Nylon [2]	147	2,570	2,717	1·09	5·23	6·33
Unstretched Nylon [2]	525	7,315	7,840	2·13	21·42	23·55

$$c_1\Delta H^{\circ}_{\text{sol}} + c_2\Delta H^{\circ}_{\text{total}} \quad . \quad . \quad . \quad . \quad (20)$$

where c_1 and c_2 are the moles of free and combined water respectively in one " mole " of polymer, and are given by the two terms of the isotherms (15).

The heat of wetting, as ordinarily defined is thus :

$$(c_{1\,\text{sat.}} - c_1)\Delta H^{\circ}_{\text{sol}} + (c_{2\,\text{sat.}} - c_2)\Delta H^{\circ}_{\text{total}} \quad . \quad . \quad . \quad (21)$$

Both these expressions must be corrected for the latent heat of vaporisation if water vapour is used as the source of water. The heat of absorption, as than the value found for M. It is suggested, then, that only a fraction of those units is accessible to the water molecule, the inaccessible fraction being given by $(M - R)/M$. This fraction presumably represents the "crystalline" portions of the fibres, so that the fraction $(M - R)/R$ may be interpreted as the so-called "crystalline to amorphous" ratio. These fractions are shown in Table V.

---

[9] Loasby, *J. Text. Inst.*, 1943, **34**, P.45.

These results assign much the same order of crystallinity to the fibres as does the X-ray evidence.

Perhaps the most serious deficiency in our model, is its failure to provide for the hysteresis which is so noticeable a feature of the experimental adsorption-desorption isotherms for the fibrous polymers, except Nylon. The suggestion is offered that the fraction of inaccessible groups is not constant throughout the absorption process, but that as absorption, and swelling, proceed some portions of the crystallites open, and the amount of hydrate formation, and possibly of solution of water also, exceed that calculated on the assumption that M is constant. The hysteresis is then due to the fact that when the saturated fibre is desorbed, there is a lag in the re-formation of the crystallites. Indeed it has not, so far as we know, been shown that the density after exposure to water, and dehydration, ever returns to its original value, or that the regain-humidity cycle can be exactly repeated.

TABLE V.

Substance.	Fraction of Groups Inaccessible.	"Crystalline to Amorphous" Ratio.
Wool [8]	0·44	0·76
Wool [4] 80's Merino	0·48	0·92
Wool [4] Oxford Down	0·44	0·79
Wool [2]	0·44	0·80
Hair [5]	0·43	0·76
Silk [6]	0·79	3·76
Silk [2]	0·80	3·91
Nylon [9]	0·85	5·67
Unstretched Nylon [2]	0·86	6·12
Stretched Nylon [2]	0·85	5·67
Cotton [7]	0·68	2·08

The free energy data of Table IIIA show some similarities. The standard free energies of solution of water in the fibres are all approximately 170 calories for wool, hair and drawn Nylon, and approximately 110 calories for silk, cotton, and undrawn Nylon. The standard free energies of hydrate formation from liquid water and dry fibre are approximately 1200 calories for wool, hair, silk and one sample of Nylon, whilst for cotton the figure is 1600 calories. The Nylons examined by Bull, however, show a lower figure of 800 calories.

The heats of reaction $\Delta H^\circ_{\text{chem}}$ and $\Delta H^\circ_{\text{total}}$ calculated from our analysis of Bull's absorption data on wool are -5416 and -6473 calories per mole and these values are very similar to the value of -5800 calories per mole calculated by Speakman from Hedges' calorimetric data.

Finally, the agreement on the one hand between the calculated specific volume-regain curve on the basis of the isothermal data and the observed curve, and on the other, the finding of the right order of the compressibility for water, must be considered as at least satisfactory.

We submit, therefore, that our model merits consideration as the basis for further experimental and theoretical studies; in particular, we suggest that the effective impossibility for phase separation to occur in the high molecular species may impose restrictions on their activity coefficients and that the excellence of fit obtained by our assumption of ideality is a reflection of this.

Imperial Chemical Industries Limited,
Research Laboratories,
Hexagon House, Blackley,
Manchester 9.

GENERAL DISCUSSION *

Dr. G. Salomon (*Delft*) (*communicated*) : The theory of van Arkel can be applied to an interpretation of swelling equilibria. Natural rubber imbibes very similar quantities of liquid polyhalides, but typical differences have been found for copolymers with a high concentration of nitrile groups. The swelling power for these elastomers decreases in the direction

$$CHCl_3 \longrightarrow CCl_4,\ CS_2 \text{ and } CHCl_2{-}CHCl_2 > CHCl{=}CHCl > CHCl{=}CCl_2 \gg CCl_2{=}CCl_2.$$

The general solvent power of hydrocarbons increases not only from aliphatic to aromatic liquids, as van Arkel states, but also from aliphatic to hydroaromatic liquids. This indicates that cyclisation contributes also to the solvent properties of an aromatic compound. The increase in intermolecular forces in liquid ring compounds can be derived from the higher boiling point of e.g., cyclohexane as compared with hexane and is probably caused by reduction of free rotation of the CH_2 groups.

Dr. G. Gee (*Welwyn Garden City*) (*communicated*) : While I am fully in agreement with van Arkel that many factors enter into the determination of the heat of mixing of two liquids, it must be pointed out that his paper does scant justice to the pioneer work of Hildebrand on the subject. The first equation quoted by van Arkel is only valid for liquids of equal molar volume. Hildebrand's general equation gives the heat of mixing per cc. of mixture as :

$$\Delta H^m = v_1 v_2 \left(\sqrt{\frac{E_1}{V_1}} - \sqrt{\frac{E_2}{V_2}} \right)^2,$$

where v_1 and v_2 are the volume fractions of the two liquids, and E_1/V_1 and E_2/V_2 are their cohesive energy densities. This latter equation is in perfect qualitative agreement with the observed immiscibility of water and heptane, and would also place the water solubilities of the chlorinated methanes in the correct order. Hildebrand has indeed had striking success in the quantitative treatment of solubility problems with this equation. Such success is only to be anticipated when dispersion forces are of predominant importance, but this is known to be true for a wide range of liquids.

I have shown that the same equation may be used for polymer liquid systems, though generally only in a qualitative way, and within a limited range of liquids. It is not surprising that Alfrey, Justice and Nelson should, in their paper, report quantitative discrepancies between theory and experiment for the systems they employ.

In spite of its defects, and its admittedly rather slender theoretical basis, Hildebrand's equation affords in my opinion by far the most valuable single method at present available to predict heats of mixing and solubilities in both liquid-liquid and liquid-polymer mixtures.

Dr. M. Magat (*Paris*) (*communicated*) : The Scatchard-Hildebrand formula usually gives only the order of magnitude of the heat of mixing. I would like to quote a case where the numerical agreement obtained by its use is unexpectedly good. In order to test how far this formula can be used in the case of polar polymers or polar solvents, we have applied it, or rather its extension for mixed solvents, as suggested by Scott and myself,[1] to the case of precipitation of cellulose acetate (Table I) and its

TABLE I.

Precipitant.	cc. added to 100 cc.	α_m.	V_m.
Methanol .	12·21	9·45	80·2
Ethanol .	25·3	9·49	80 1
n-Propanol .	33·0	9·48	80·6
n-Butanol .	40·5	9·47	81·0

* On two preceding papers.

[1] Scott and Magat, *J. Chem. Physics*, 1945, 13, 172.

distribution between two solvents (Table II). The experimental data are taken in both cases from the work of Dobry,[2] where ψ_i = volume fraction of the liquid i, and V_i = molecular volume of the liquid i.

$$\alpha_m = \psi_1\alpha_1 + \psi_2\alpha_2 ; \qquad V_m = \frac{V_1V_2}{\psi_1V_2 + \psi_1V_2}$$

$$\alpha = \sqrt{\frac{E}{V}}.$$

The agreement here is quite striking.

The distribution of the cellulose acetate between the solution and the precipitate[3] (coacervate) is calculated according to Scott and Magat (*loc. cit.*) with the formula

$$\ln \frac{\psi_p'}{\psi_p} = m_i \left\{ 2\mu(\psi_p' - \psi_p) - \ln \frac{1 - \psi_p}{1 - \psi_p'} \right\}.$$

The value of α = 10 for cellulose acetate was deduced from the osmotic pressure slopes. The molecular weight was assumed to be equal to the average molecular weight of the fraction, 100,000. No further adjustments were undertaken, $1/\gamma$ being assumed equal 0·28 throughout all our calculations. A more complete account will be published soon in *J. Chim. Physique.*

TABLE II

ψ_p Coacervate.	ψ_p Solution.	μ (Calc.).	Per cent. Alcohol in the Alcohol-Chloroform Mixture. (Calc.).	(Obs.).
6·6	0·4	0·545	3·3	4·4
6·35	0·45	0·540	3·7	4·7
5·5	0·8	0·538	3·8	5·3
5·3	1·2	0·537	3·9	5·9
5·5	1·6	0·536	4·0	6·4
4·3	2·2	0·530	4·4	6·6

Dr. P. H. Hermans (*Utrecht*) said: The assumption of Hailwood and Horrobin that the formation of true hydrates plays a part in polymer-water systems showing isotherms of a sigmoid shape is substantiated by our own work. In the case of the cellulose-water system the formation of true hydrates can be derived from X-ray data,[4] a fact which has been as yet overlooked. We have also given a similar theoretical interpretation of the isotherm.[5] It may be of interest to the authors that we have recently published accurate data on the specific volume of regenerated cellulose as a function of regain[6] which may be used to check their theory besides those of King for wool. The crystalline amorphous ratio found for cotton by Hailwood and Horrobin is of the same order of magnitude as that derived by us from density and other physical data.[7]

There is perhaps one objection to be made to the paper. It would seem that the considerations regarding the compression of the water have no real physical meaning. The volume contraction occurring at swelling is—in my opinion—of the same nature as that occurring upon dissolution of sugar in water or in the system, sulphuric acid-water. There is no question of a compression of the water (or of the other component) and it is even doubtful whether compression is a suitable concept in this connection. As we have set forth elsewhere, the phenomenon could just as well be explained by purely geometrical considerations and be compared with the changes in volume occurring on the mixing of spheres of different radius.[8]

---

[2] *J. Chim. Physique*, 1938, **35**, 387.

[3] Dobry, *ibid.*, 1938, **35**, 387.

[4] Hermans and Weidinger, *J. Colloid Sci.*, 1946, **1**, 185.

[5] Hermans, *Contribution to the Physics of Cellulose Fibres* (Elsevier, Amsterdam–New York, 1946), p. 41, 188.

[6] *Ibid.*, p. 73.

[7] Hermans and Weidinger, *J. Amer. Chem. Soc.* (in press).

Dr. Cassie (*Leeds*) said : Hailwood and Horrobin have worked out an isotherm for the sorption of liquid by polymers using the solution hypothesis applied to essentially the same model as that used in the adsorption theory. It is, therefore, of interest to compare the results given by the two theories with experimental data.

The authors state that adsorption theories have had less success than their own, but it must be remembered that the adsorption equations to which they refer, apply only to free surfaces, and not to adsorption into the bulk of an elastic gel. I have modified the adsorption equations of Brunauer, Emmett and Teller [9, 10] to allow for adsorption into the bulk of a gel, and find excellent agreement with the empirical isotherm for water absorbed by keratin.

Hailwood and Horrobin state that they assume that dissolved water and hydrated and unhydrated polymer form an ideal solution, but it is clear from equation (3) of their paper that the isotherm equation is not obtained from this assumption alone. Equation (3) really states that the dissolved water is not ordinary liquid water, and equation (24) reveals that they assume the difference between liquid and dissolved water to be due to compression. They have, in fact, corrected their simple solution equations in qualitatively the same way as I have corrected the adsorption equations. If this modification is omitted, the regain becomes infinite at 100 % R.H., as it should do for an ideal solution, and as may be seen from equation (14) when α is unity. The agreement with the observed isotherms is then even poorer than that given by Brunauer, Emmett and Teller's unmodified equation.

The solution theory is verified largely with isotherm data, but as the isotherm equation, (15), has three disposable parameters, this agreement gives little check on the correctness of the theory. The modified adsorption theory gives equally good agreement with isotherm data, but again with virtually three disposable parameters. Other data must, therefore, be sought to decide which is the better starting hypothesis.

Heat data undoubtedly supply the best criteria for a theory whose constants have been deduced from empirical isotherms. Hailwood and Horrobin state that the agreement of their calculated heats of wetting with Hedge's empirical data " is only fairly good at the lower humidities," and dismiss the poor agreement as due to errors in Hedge's data. Hedge's data are much abused, but the heats of wetting of wool fibres, powdered wool, and powdered horn keratin have been repeatedly measured in the Wool Industries Research Association's laboratories, and in all cases Hedge's curve of heat of complete wetting against regain has been verified ; in fact, they are probably the most reproducible data available for the keratin-water system. Hedge's data are available only at regains below 18 % ; above this regain, nothing is known about heats of wetting apart from the very significant curve of regain-temperature obtained by Speakman, Stott and Chang.[11] At constant humidity near saturation vapour pressure, this curve shows a minimum at 37° C. Any reasonable theory should reproduce this minimum, and it may be significant that the modified adsorption theory does so.[8] Finally, a useful theory should explain the diffusion data given by King for water in keratin.[12] Here, again, the adsorption theory appears to be successful.

---

[8] Ref. [3], p. 58.

[9] Cassie, *Trans. Faraday Soc.*, 1945, **41**, 458.

[10] Cassie, *Soc. Dyers Col., Symposium on Fibrous Proteins*, 1946.

[11] Speakman, Stott and Chang, *J. Textile Inst.*, 1933, **24**, T 284.

[12] King, *Trans. Faraday Soc.*, 1945, **41**, 479.

Dr. W. W. Barkas (*Princes Risborough*) said: There is a difficulty in evolving theoretical expressions for the sorption isotherms of solid polymers in that the process of sorption, if accompanied by swelling, must set up stresses in the polymer network which will tend to prevent further sorption. This is thermodynamically inevitable because the degree of sorption in a hygroscopic material depends both on the surrounding vapour pressure and on the magnitude of any applied stresses. Thus the calculated moisture contents which might well be correct for a stress-free system, will be either: (*a*) too high at high vapour pressures if the polymer network is stable but the stresses increase with swelling, or (*b*) too low if the polymer network becomes more easily deformable by the process of sorption, so that the stresses decrease with the degree of swelling.

In Table I of the authors' paper are several fibres all of which soften under swelling, and in all the calculated sorptions at high vapour pressures are somewhat lower than the observed, thus following case (*b*).

As regards sorption hysteresis it can be shown that in the swelling of a rigid material or molecular network, any mechanical restraint which opposes the swelling and, at the same time, induces shear strains therein, must lead to hysteresis in the isotherm unless the material is perfectly elastic. (The converse, that the sorption hysteresis can only be caused by mechanical restraints in swelling, does not necessarily follow.) It is interesting to note that the authors quote Nylon as a material in which hysteresis is small. A synthetic Nylon fibre, even if extruded under tension, has a simpler structure than any natural fibre and is less liable to develop internal shear strains during swelling.

Dr. G. Gee (*Welwyn Garden City*) (*communicated*): The quantitative treatment presented in this paper * seems to me to rest on a completely false premise. The solid solution of wool, hydrate and water is assumed ideal, whereas an outstanding thermodynamic characteristic of polymer solutions is the extent to which they deviate from ideality. This deviation may be represented by calculating, from the vapour pressure, an " apparent molecular weight " M', assuming Raoult's law, and noting its enormous dependence on concentration. Thus for rubber in benzene, M' varies between 300,000 and 300 as the mole fraction of rubber increased from 0 to 1. I suspect that what the authors of this paper have in fact done is just this. If so, the molecular weights they obtain are mathematical fictions, and bear no relationship whatsoever to any unit in the polymer structure.

Dr. R. M. Barrer (*London*) (*communicated*): The importance both technological and theoretical attaching to the problem of the sigmoid isotherm in polymers makes each attempt at interpretation of significance. The authors' model and results are therefore of considerable interest. It is necessary, however, to examine the assumptions on which the isotherm is based with care for it must be remembered that any three constant equation of algebraic form

$$A + Bx - Cx^2 = \frac{x}{7},$$

must for a suitable choice of constants follow an S-shaped course. To what extent therefore is agreement with sigmoid isotherms empirical and to what extent has it physical significance? To answer this question the underlying assumptions must then be critically examined.

One of the assumptions cannot it seems to the writer be correct. This is that the three species present in the solid phase (water, polymer hydrate and polymer) can form an ideal solid solution, because even if the heat of mixing were zero, one of the molecules (water) is small and the others very large. Statistical thermodynamic consideration of the mixing of large with small molecules reveals marked departures from the ideal mixture, and these departures become greater the larger the discrepancy in size between the species being mixed.[13] Thus, for the polymer systems being considered by the authors, the properties of the mixture will show the maximum departure from the ideal.

---

* By Hailwood and Horrobin.

For the case of simple solute molecules and of long flexible randomly coiling chain molecules, the statistical thermodynamics of mixing has been considered both for zero [14] and non-zero [15] heats of mixing, and this treatment has been extended to cross-linked networks of flexible chains imbibing liquids.[16] It has also been shown how the isotherm (vapour pressure of solute-composition of mixture) follows both above and below the critical temperature of the solute.[17] While there are still imperfections in the theory it has been shown to be a close approach to the behaviour of rubber-like materials in presence of liquids,[18] and one is led to enquire whether the S-shaped sorption isotherm in other chain polymers cannot be approached in a similar way.

In a recent discussion the writer [19] indicated how the combination of the process of mixing of small molecules and long chain molecules coupled with a simultaneous and independent sorption of Langmuir type on localised sorption sites distributed through the polymer, leads naturally to an S-shaped isotherm. This localised sorption corresponds to uptake of water at polar groups in the structure (CO, NH, etc.) or by polar impurities.

It has to be recalled that while the model upon which this isotherm is based is an improvement upon the assumption of an ideal solid solution, it can only be qualitative at present for many polymers, where much of the material is crystalline or where the polymer molecules are not fully flexible. There may also be further difficulties in the assumption of an ideal solid solution made by Hailwood and Horrobin. However, the present comments suggest the way in which an adequate interpretation of some S-shaped isotherms may eventually be evolved, an interpretation at present applicable to elastomers containing polar impurities. Later developments of the statistical theory of mixing may cover the cases of small molecules admixing with crystalline chain polymers, or comparatively inflexible polymers. The isotherm should be nearly valid for those proteins referred to in this Discussion [20] in which sorption of water and other molecules reduce lateral attraction between chains sufficiently to cause random coiling of chains and rubber-like elasticity of the mixture.

Dr. Vermaas (*Utrecht*) (*communicated*) : The following may serve as an illustration of Dr. P. H. Hermans's remarks, bringing further details concerning our own work on the water sorption of regenerated cellulose and a new evaluation of King's density data for moist wool.

In their interesting paper dealing with the sorption of water by polymers the authors demonstrate that their formula (16) gives a good description of the sorption isotherm of various materials in the whole region of regain from 0-100 %. To the examples given in Tables I and II, I should like to add that of isotropic regenerated cellulose which has been studied extensively in our laboratory.[21] It is shown in Fig. 1 how in this case, too, the water sorption as a function of regain is described excellently by formula (16), using the constants $A = 1{\cdot}43$, $B = 0{\cdot}103$ and $C = 0{\cdot}00088$. This implies a value $M = 219$ and a percentage of inaccessible cellulose of 35 % ($R = 162$). The latter value is in good agreement with the percentage of crystalline material as computed from density determinations,[22] heats of wetting and the X-ray diffraction of amorphous cellulose.[23] In order to obtain a good agreement between the experimental regain values and the theoretical curve, it suffices to assume the formation of a monohydrate of cellulose only. There are, however, good reasons to assume the existence of two distinct hydrates in cellulose, viz.

---

[13] e g. Fowler and Guggenheim, *Statistical Thermodynamics* (C.U.P., 1939), p. 360.

[14] Flory, *J. Chem. Physics*, 1942, **10**, 51; Huggins, *Ann. N.Y. Acad. Sci.*, 1942, **43**, 1, Miller, *Proc. Camb. Phil. Soc.*, 1942, **38**, 109; 1943, **39**, 54, 131.

[15] Orr, *Trans. Faraday Soc.*, 1944, **40**, 320.

[16] Flory and Rehner, *J. Chem. Physics*, 1943, **11**, 512, 521; Flory, *Chem. Rev.*, 1944, **35**, 52.

[17] Barrer, *Trans. Faraday Soc.* (in press).

[18] Gee and Treloar, *ibid.*, 1942, **38**, 147; Gee and Orr, *ibid*, 1946, **42**, 507; Gee, this Discussion.

[19] Barrer, *Soc. Dyers Col.*, *Symposium on Fibrous Proteins*, 1946.

[20] Jordan-Lloyd, Dempsey and Garrod, this Discussion.

[21] Hermans, *Contribution to the Physics of Cellulose Fibres* (Elsevier's Publishing Co., Amsterdam–New York, 1946), p. 13.

$$C_6H_{10}O_5 \,.\, 1/3\ H_2O \text{ and } C_6H_{10}O_5\ 4/3\ H_2O$$

respectively.[24] The assumption of two hydrates involves the introduction of a further constant into the sorption equation, allowing of a still closer approach of the theoretical curve to the experimental regain values.

We have also investigated the dependence of the specific volume on regain for regenerated cellulose. The resulting curve strikingly resembles that for moist wool as determined by King. It is, therefore, tempting to interpret both curves in an analogous way. Our method of interpretation differs from Hailwood and Horrobin's in some respect, using a graph as shown in Fig. 2. Here the volume of 1 g. of wool after the uptake of $r/100$ g. of water has been plotted against $r/100$. The curve starts with a slope smaller than 45° at low regains, but beyond about 20 % regain it exhibits a nearly exact 45° slope. The slope at a certain regain r represents the apparent volume of 1 g. of water absorbed by an infinite mass of moist wool of regain r.

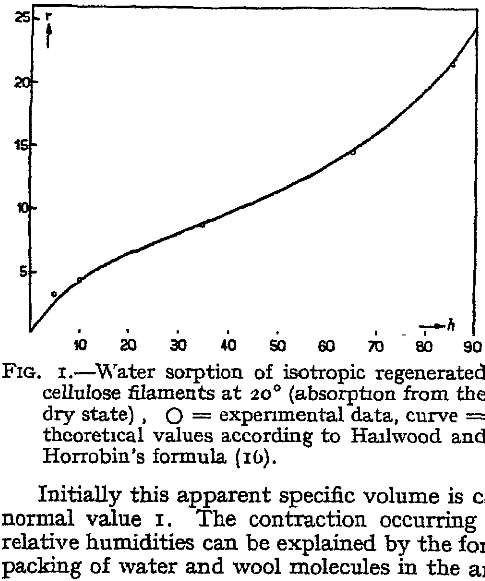


FIG. 1.—Water sorption of isotropic regenerated cellulose filaments at 20° (absorption from the dry state), ○ = experimental data, curve = theoretical values according to Hailwood and Horrobin's formula (16).

Initially this apparent specific volume is considerably smaller than the normal value 1. The contraction occurring on taking up water at low relative humidities can be explained by the formation of a more favourable packing of water and wool molecules in the amorphous parts of the fibre; on sorption of water by dry or nearly dry wool a certain amount of "empty space" disappears. Beyond about 20 % regain the water is absorbed with its normal density. Here all "empty space" has been filled and packing is as dense as possible.

The hypothesis of a real compression of the water is entirely superfluous. It is possible to derive from the graph the volumes of "empty space" (a), water (b) and wool (c) at any regain. It should be noted that the quantity c here represents the volume of 1 g. of wool molecules when packed with water molecules as densely as permitted by the structure of the swollen wool fibre. In other words, c must be the specific volume of wool measured in water. In fact the point of intersection of the extrapolated linear part of the specific volume curve and the ordinate in Fig. 2 coincides exactly with the specific volume of wool in water as determined by King (point B).

---

[22] Hermans, *loc. cit.*

[23] Hermans and Weidinger, *J. Amer. Chem. Soc.*, 1946, 68, 1138.

[24] Hermans and Weidinger, *J. Colloid Sci.*, 1946, 1, 185.

It will be clear that this explanation holds equally well if hydrates are formed at sorption. On the other hand it demonstrates the cause of the failure of Hailwood and Horrobin's formula (25) at high regains. It is not justified to assume a constant value of 15·32 cc./mole for the apparent molecular volume of water in moist wool in the whole region from 0 to 100 % regain. From Fig. 2 one derives a limiting value of 10·9 cc./mole at 0 % regain, whereas at high regain the differential molecular volume of the sorbed water approaches the normal value 18. In formula (25), v_1 represents an average molecular volume. The value 15·32 given by the authors applies reasonably well in the region of low regains but beyond $r = 20$, the systematic deviations of King's values from the authors' curve are an indication of a gradual increase in v_1. In order to describe the point of 33 % regain in Fig. 1 of Hailwood and Horrobin's paper by their formula (25) one has to substitute $v_1 =$ 16·6 cc./mole, taking for v_2 and v_3 the values 179·1 and 168·1 respectively.

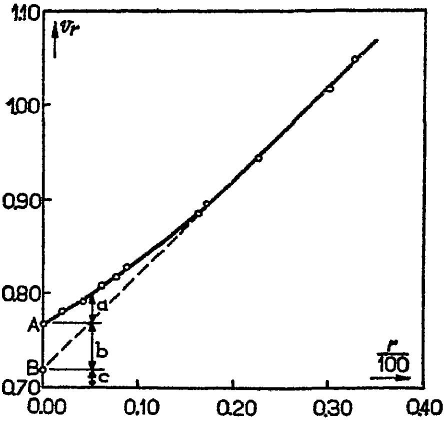


FIG. 2.—Vol. of 1 g of wool after sorption of $r/100$ g. of water/regain (King's data). A = specific vol. of dry wool in benzene, B = specific vol. of wool in water, a = "empty space," b = vol. of sorbed water, c = vol. of wool.

Prof. A. R. Ubbelohde (*Belfast*) said: Statistical theories about polymers usually assume that complete equilibrium has been established with respect to all possible configurations of the molecules. It seems desirable to review the essential conditions for thermodynamic equilibrium, before attempting any final evaluation of experiments designed to test such statistical theories.

The test of true thermodynamic equilibrium is that small fluctuations within the system do not lead to any trend. This test of thermodynamic equilibrium in terms of the effect of small fluctuations bears a formal analogy to the principle of virtual work in testing mechanical equilibrium. At equilibrium, in systems at constant pressure the free energy G is a minimum, and $\pm\delta G = 0$ for all possible small changes in the system. A further condition is that these small changes must be possible kinetically, which can only happen when the fluctuations involved are of sufficient frequency to permit the adjustment of the system to a condition of minimum free energy. The relaxation time for the fluctuations comprised in the condition $\pm\ \delta G = 0$ must be small compared with the time of the experiment.

These two conditions for complete equilibrium may be illustrated with reference to well-established cases where the equilibrium is incomplete. In such cases of incomplete or partial equilibrium, the criterion $\pm\ \delta G = 0$ is kinetically significant with respect to certain fluctuations, but with respect to others it has no experimental significance. The steady state reached with respect to these changes for which the relaxation time is large may be far removed from complete equilibrium, and may be in no sort of agreement with statistical theories which assume complete equilibrium with respect to all possible configurations of the system.

(i) *Partial equilibrium with reference to changes of atomic linkages:* to select only one out of a very large class of examples, a mixture of H_2 and O_2 at ordinary temperatures reaches thermal equilibrium with reference to the translational energy of the molecules, in a very small relaxation time. Owing to restrictions on the transfer of rotational energy (including the *ortho-* and *para-*states of the hydrogen molecule) the relaxation time for rotational equilibrium is probably somewhat larger, but is still small compared with most experiments. On the other hand, the kinetics of the reaction,

$$2H_2 + O_2 \rightleftharpoons 2H_2O$$

are so unfavourable at room temperatures, in the absence of catalysts, that fluctuations of free energy which involve an exchange of atomic linkages are normally of no experimental significance. The system does not reach minimum free energy with respect to this type of change.

(ii) *Partial equilibrium with reference to changes of configurational entropy:* an example of partial equilibrium closely similar to that of many polymer systems is presented by a liquid supercooled below its freezing point. Normally a supercooled liquid is in thermal equilibrium with respect to the distribution of translational and rotational energy. It can also be in equilibrium with respect to the transition, liquid ⇌ vapour. But the kinetics of the reaction, liquid → solid, are so unfavourable in the absence of nuclei, that the fluctuations of free energy which involve this transition are often not realised experimentally. In this instance, the changes of configurational entropy sufficient to produce a " nucleus " of solid from the liquid, require the co-operation of an appreciable number of molecules. Kinetically, such co-operative processes are very improbable.[25] In view of this feature, it seems likely that whenever large changes of *configurational entropy* are required to reach equilibrium, the relaxation time may be so large compared with the time of experiment, as to make the condition $\pm \delta G = 0$ meaningless with respect to such changes.

The experimental fact that in systems in partial equilibrium the equilibrium is quite definite with respect to certain changes, and meaningless with respect to others, implies that statistical theory can only be applied to such systems, with respect to carefully defined changes. The usual procedure for calculating the maximum probability of distribution of molecules in different energy states is experimentally relevant only when the corresponding relaxation time is sufficiently small. In such cases, it may still be possible to formulate a satisfactory partition function, by treating those factors which refer to irrelevant changes as arbitrary constants.

Experimental evaluation of the relaxation time for equilibrium in polymer systems: Considerable experimental ingenuity may be called for in verifying how far a steady state in a polymer system can in fact be reached from both sides, i.e. is a state of complete equilibrium. Two experimental procedures may be quoted to illustrate the sort of information required for the particular case of the uncoiling or uncrumpling of a long chain molecule. When the structure of the coil is such that the innermost parts have no freedom to unwind until the outermost parts have moved out of the way, the succession of steps from a coiled to a completely uncoiled molecule is kinetically analogous to a chain reaction. The probability of spontaneous occurrence of such a connected sequence of events may be very different from the probability of a series of independent events, and it is important to obtain experimental evidence about the rate at which the various types of polymer molecule coil and uncoil.

(i) If the polymer is subjected to static tension, the molecules will uncoil. If an oscillating tension is applied, then when the frequency of oscillation is small the molecules will coil and uncoil much as under static tension. But as the frequency of oscillation is increased, and becomes comparable with the time required for coiling and uncoiling, the elasticity of the molecule may be expected to show a marked drop. The critical frequency of oscillation for this change of elasticity may be expected to give a good idea of the order of time required for coiling and uncoiling. Detailed statistical analysis is somewhat complicated in this instance by the mutual interference of the molecules, but the general features of the statistical analysis are closely similar to those for other " relaxation " phenomena, such as the decrease of bulk modulus of CO_2 with increasing frequency,[26] or dielectric relaxation effects.

---

[25] Ubbelohde, *Trans. Faraday Soc.*, 1937, 33, 1203.

Important information about the differences between different polymer molecules should be obtainable from a systematic investigation of what activation energy, if any, is associated with the uncoiling reaction, particularly in cases where fresh cross links are formed. Similarly, the effect of absorption of water and other solvents on the critical frequency may give useful information on the nature of the processes of absorption, as well as testing the completeness of thermodynamic equilibrium.

(ii) A study of the coiling and uncoiling of polymer molecules isolated from their neighbours is somewhat simpler theoretically, and more difficult experimentally. It may be possible, for example, to squirt a solution of the polymer in a bad solvent, into a large excess of " good " solvent, and vice versa, and to follow the rate of coiling or uncoiling of the polymer molecules from the light scattering.

These two illustrations of the sort of experimental investigation required to evaluate relaxation times are not in any way exhaustive, but are meant to emphasise the point that thermodynamic ideas about equilibrium should not be applied recklessly to polymer systems, until the appropriate relaxation times have received experimental study.

Dr. L. R. G. Treloar (*Welwyn Garden City*) said : Prof. Ubbelohde is entirely correct in his remarks concerning the necessity of ensuring that configurational changes in high-polymer molecules are kinetically realisable before applying the methods of statistical thermodynamics. In rubbers, however, there are many independent lines of evidence which show that the theory is in fact applicable, at least to a first approximation. There are, firstly, direct measurements of relaxation times, among which may be quoted Aleksandrov and Lazurkin's data [27] on rubber in vibration, which yield relaxation times of the order 10^{-6} sec. at room temperature, and secondly, studies of many physical properties (elasticity, solution viscosity, swelling phenomena, etc.) in which the predictions of the statistical theory are substantially verified. In a recent work of Gee,[28] for example, the question of obtaining a genuine equilibrium is carefully considered, and the free energy change accompanying the extension of rubber is analysed and compared with the theoretical value.

It might be appropriate here to draw attention to the unsuitability of the word *coil* to describe the form of a long-chain molecule of the rubbery type. An essential feature of a *coil* is regularity ; the essential feature of the molecule is lack of regularity. I prefer to use the term *randomly-kinked chain*, and to refer to *unkinking* rather than uncoiling. These terms are admittedly clumsy, and any suggestion for an improvement in this nomenclature would be welcome.

Messrs. Hailwood and Horrobin (*Manchester*) (*communicated*) : We note with interest that Dr. P. H. Hermans and his co-workers have obtained X-ray and other evidence of the presence of hydrates in moist cellulose, and that they had been treating the problem of the adsorption isotherm on similar lines.

We welcome the additional example given by Dr. Vermaas showing the fit of our isotherm to their data on regenerated cellulose, and pointing out that the crystalline/amorphous ratio deduced in this case accords well with estimates by independent measurements.

[26] Various authors, e.g. Richards and Reid, *J. Chem. Physics*, 1934, **2**, 194
[27] *J. Tech. Phys. U.S.S.R.*, 1939, **9**, 1249.
[28] *Trans. Faraday Soc.*, 1946, **42**, 585.

The concept of a pressure on the dissolved water should not be taken too literally. It would perhaps suffice to say that this water is in a different field of force from that in ordinary liquid water.

We are at present studying Dr. P. H. Hermans' monograph, to which he has referred, as well as the additional data given by Dr. Vermaas.

Dr. Cassie suggests that in deducing our isotherm from our model, we have assumed not only (1) that the mixture of our three species is ideal, but also (2) that the dissolved water is not ordinary liquid water, and that we further assumed (3) that the difference is due to compression. In fact, (2) is a consequence of (1) and is not an independent assumption. It is unnecessary therefore to assume any mechanism such as (3) to account for it, and certainly we did not arrive at (2) by way of (3). We agree with Dr. Cassie that heat data will provide a valuable check on any isotherm whose constants have been deduced by fitting to the experimental curves. It is essential, however, that all the data should be obtained on the same specimen of wool or other fibre and we should welcome the publication of such a set of results which would enable this check to be made.

The considerations advanced by Dr. Barkas seem to us important, and to express much more precisely than we had done the ideas as to the origin of the hysteresis which we had had in mind towards the end of our paper.

Dr. Gee seems to have misunderstood our assumption of ideality. The mixture of polymer and water regarded as a two-species system is of course far from ideal, as indeed are many simpler two component mixtures such as, e.g. water and alcohol. We have shown that this apparent non-ideality can be accounted for if the system is regarded as an ideal solution of the anyhdrous polymer, polymer hydrate, and water (i.e. of three species). It likewise appears that he has not understood our mode of calculating the weight of the operative molecular unit, which we do not in fact assume to vary with regain.

Dr. Barrer's remarks concerning the three constant equation of the given algebraic form are certainly true, but it must be realised that we did not choose this form empirically with a view to fitting sigmoid isotherms, but have reached it as a result of our theory.

He rejects our assumption of ideality on the ground of the statistical analysis of the case of two types of molecules of markedly different size which is presented by Fowler and Guggenheim. This analysis, however, applies to the cases where the different molecules are interchangeable in position. This is not the case where water molecules are free to wander within the spaces of a polymer structure which is substantially immobile, and we have suggested that this is in fact a case in which the mixture could not be other than ideal. We agree that the point is one of fundamental importance and merits exhaustive theoretical analysis.

ON THE INTERACTION OF COLLOIDAL PARTICLES. I. PARTICULAR APPLICATION TO PARALLEL PLATES.

By S. Levine.

Received 13*th August*, 1946.

As revised 28*th October*, 1946.

1. Introduction.

The plate-like particles in sols of Fe_2O_3 or WO_3 arrange themselves in parallel, horizontal layers (Schiller layers) on the bottom of the vessel which contains the sol, with distances of separation between the layers as great as 8000 A.[1] Similarly, soap [2] and lipide [3] micelles, which consist of flat plates, are also separated by water layers reaching thicknesses, however, of the order of only 100 A. It might be assumed in the former case that equilibrium exists between the repulsive forces between the particles associated with their overlapping double layers and the attractive force of gravity. However, with the micelles the attractive force necessary to produce the equilibrium distances observed must have a different origin, and furthermore, the much smaller spacing suggests that different factors come into play.

The existence of long range attractive forces between colloidal particles in an aqueous dispersion medium, changing to repulsive forces at smaller interparticle distances (and thus having a position of equilibrium) is demonstrated by the formation of tactoids in sols of V_2O_5, benzopurpurin,[4] and Fe_2O_3,[5] and in solutions of tobacco mosaic virus protein,[6] and of a fibrous modification of insulin protein.[7] All these systems are thixotropic. In the case of the virus protein, Bernal and Fankuchen [6] have shown by X-ray analysis that the tactoids consist of parallel and equidistant rod-like particles forming a hexagonal lattice in cross-section. The distances between the particles can be as much as 500 A. Thixotropic gels of plate-like bentonite [8] do not manifest the tactoid formation, but the fact that the gels can be obtained at concentrations as low as 0·05 % again suggests the presence of long-range forces [9]

Two points of view have been expressed on the nature of the far-reaching forces between the colloidal particles. One is that the repulsion arises from the interaction of the electrical double layers of the particles, whereas the attraction comes from the van der Waals-London dispersion forces. This was first proposed by Kallmann and Willstatter [10] and developed by Freundlich [11] and Hamaker.[12] Recently, Verwey [13] has supported this theory. The other viewpoint has been advanced by Langmuir,[8] who considered primarily the bentonite plates, and by the author who, in collaboration with Dube, treated spherical [14] and cylindrical [15] particles. Langmuir [8] considered that both the attractive and repulsive forces between the particles were electrical in nature. However, we agree with Verwey's [13] criticism that Langmuir's general theory is untenable. Dube and the author calculated the mutual electrostatic interaction of two spherical colloidal particles and obtained a minimum in the potential energy. The assumption was made that the charge on the particles did not change with interparticle separation. This work was criticised by Derjaguin[16] and Verwey[13] on the grounds that the free energy and not the electrostatic potential energy should be considered. Furthermore, both Derjaguin and Verwey assumed that the electrical potential at the surface of the particles is constant, independent of the particle separation. This means that the number of ions adsorbed on the particle surface from the solution changes with the separation, and Verwey introduced the " chemical " part of the adsorption energy. The conclusions of Derjaguin and Verwey were that the interaction of the double layers of the colloidal particles yields a repulsion and that any attraction is to be attributed to the van der Waals-London forces.

---

[1] Zocher, *Z. anorg. Chem.*, 1925, **147**, 91 ; Zocher and Heller, *ibid*, 1930, **186**, 75 ; Heller, *Compt. Rend.*, 1935, **201**, 831 ; Bergmann, Low-Beer and Zocher, *Z. physik. Chem. A*, 1938, **181**, 301.

[2] Harkins, Mattoon and Corrin, *J. Amer Chem. Soc.*, 1946, **68**, 220, where references to the earlier German work are given. Also Ross and McBain, *ibid*, 1946, **68**, 296.

[3] Palmer and Schmitt, *J. Cell. Comp. Anatomy*, 1941, **17**, 385.

[4] Zocher and Jacobsohn, *Koll-Beih.*, 1929, **28**, 167.

[5] Coper and Freundlich, *Trans. Faraday Soc.*, 1937, **33**, 348 ; Heller *et al.*, in a series of papers in *Compt. Rend.*, 1936-39, and in *J. Physic. Chem.*, 1941, **45**, 1203 ; 1942, **46**, 765, 783.

[5] Coper and Freundlich, *Trans. Faraday Soc.*, 1937, **33**, 348.

[6] Bernal and Fankuchen, *J. Gen. Physiol.*, 1941, **25**, 111.

[7] Waugh, *J. Amer. Chem. Soc.*, 1946, **68**, 247.

[8] Hauser, *Chem. Rev.*, 1945, **37**, 287 ; Langmuir, *J. Chem. Physics*, 1938, **6**, 873.

[9] This conclusion has been questioned by Lawrence, *Ann. Reports*, 1940, **37**, 109, who proposes the random-mesh theory of gel structure. However, it is difficult to explain the results of Bernal and Fankuchen on the tobacco mosaic virus proteins in this way.

[10] Kallmann and Willstatter, *Naturwiss*, 1932, **20**, 952.

[11] Freundlich, *Thixotropy* (Paris, 1935).

[12] Hamaker, *Physica*, 1937, **4**, 1058 ; *Rec. trav. Chim.*, 1937, **56**, 1, 727 ; 1938, **57**, 61.

[13] Verwey, *Philips Research Reports*, 1945, **1**, 33.

[14] Levine and Dube, *Proc. Roy. Soc. A*, 1939, **70**, 145 ; *Trans. Faraday Soc.*, 1939, **35**, 1125 ; 1940, **36**, 215, 725 ; *Phil. Mag.*, 1940, **29**, 105 ; *J. Chem. Physics*, 1939 ; **7**, 831 ; *J. Physic. Chem.*, 1942, **46**, 239.

[15] Dube, *Ind. J. Physics*, 1943, **17**, 189. This work was begun in collaboration in 1939, but its joint completion was prevented by the outbreak of the war.

The author has independently derived an expression for the " electrical " part of the free energy of interaction of two colloidal particles, which is substantially the same as that used by Verwey. However, the manner in which the " chemical " term has been introduced by the latter is subject to criticism. This report is a preliminary account of an alternative theory of interaction of colloidal particles, with particular application to two parallel plates. The full mathematical details have been omitted and will be published later. The general conclusion of the following work is that the existence of a minimum in the mutual free energy associated with the double layers of colloidal particles is quite possible, without the introduction of the van der Waals forces.

The tentative calculations presented here yield such a minimum, but in the case of parallel plates a difficulty arises, namely the position of this minimum, is at much smaller distances than are found with Schiller layers.[1] Although the thickness of the water layers between the lamellar soap micelles[2] and the bimolecular leaflets of lipide systems[3] are also much smaller, our theory in its present form does not apply to these concentrated systems.

Verwey's analysis is self-consistent, but in deriving his general formula for the free energy, he assumes that the mean potential at an ion adsorbed on the surface, due to the surrounding ions, is independent of the particle separation. He also does not clearly distinguish between this mean potential and the potential of the mean force acting on the adsorbed ion. The latter approximation is similar to that made in the Poisson-Boltzmann equation which is used in determining the interaction energy. Our analysis indicates that both the mean potential and the potential of the mean force at an adsorbed ion depend on the distance. Now the mechanical strength of a thioxtropic gel is very low and hence the forces responsible for the tactoid formation are quite small. Since Verwey is strictly not calculating the mutual free energy of the colloidal particles, the slight discrepancy may be sufficient to mask the nature of the forces at large separations. In this connection, we are able to derive a general theorem relating to the free energy of interaction, which is not based on the Poisson-Boltzmann equation, and which is inconsistent with Verwey's results.

Among the possible reasons for the discrepancy between our theory and the results for Schiller layers, three perhaps are significant, namely, the use of the Poisson-Boltzmann equation, and the neglect of the finite size of the plates and of the image forces. The last effect is probably quite appreciable with the soap and lipide micelles because of the much smaller spacings between the plates.

In section 2 of this paper, a general expression for the free energy of interaction of two parallel plates is developed by applying the methods of statistical mechanics. An explicit formula for this expression can be obtained if the mean potential at an ion and the ionic densities in the double layers are determined. In section 3, an approximate method of calculating these quantities for small surface potentials, based on the Debye-Huckel approximation to the Poisson-Boltzmann equation is given. The results are compared with those of Verwey.

---

[16] Derjaguin, *Acta Physicochim.*, 1939, 10, 333 ; *Trans. Faraday Soc.*, 1940, 36, 203 ; Derjaguin and Landau, *Acta Physicochim.*, 1941, 14, 633.

2. General Formula for the Free Energy of Interaction of Two Parallel Plate-like Particles.

Consider two parallel plate-like particles at distance R apart immersed in an electrolyte solution. The area A of the plates is assumed very large compared with R^2 so that the edge effects can be neglected. The total volume V of the solution is assumed to be large compared with the volume $v = AR$ of the region between the two plates. Suppose that the surfaces of the particles have become charged by preferential adsorption of potential-determining ions of one species only, say species 1. The origin of the charge on the particles may be, of course, the ionisation of dissociable groups on the surface, but this would not affect the argument developed below.

Suppose that all the ions, i.e. those present as solute ions in the aqueous dispersion medium and those adsorbed on the surface of the particles, are completely discharged. We now carry through the fictitious process of charging all the ions uniformly, as introduced by Debye and Hückel in their theory of strong electrolytes. The fraction λ will represent the stage in charge, *i.e.* if e_i is the normal charge on ion of type i, then λe_i will be its charge at stage λ. Let $E(\lambda)$ be the total potential electrostatic energy of the ions at stage λ for a given instantaneous configuration of the ions. This can be written as the sum of two terms, namely, $E(\lambda) = E_c(\lambda) + E_s(\lambda)$ where $E_c(\lambda)$ is the interionic Coulomb energy, proportional to λ^2, and $E_s(\lambda)$ is the electrical self-energy of all the ions. Then the increase in the Helmholtz free energy of the whole system, associated with an increment $d\lambda$ during the charging process, can be expressed as

$$\frac{\partial E}{\partial \lambda}d\lambda = \frac{2}{\lambda}\overline{E}_c(\lambda)d\lambda + \frac{\partial E_s(\lambda)}{\partial \lambda}d\lambda, \quad . \quad . \quad . \quad (1)$$

where the bar indicates averaging over all possible configurations of the ions. This general result, which follows from statistical mechanics, has been discussed by Fowler,[17] Onsager[18] and Fuoss.[19] The charging process is interpreted in the following way. For a system at constant temperature T and volume, it is assumed that during each successive stage of charging, λ to $\lambda + d\lambda$, the assembly of ions is in the equilibrium state corresponding to the stage λ. There is no external work done in the rearrangement of the ions into their equilibrium state in between the stages. Thus all the ions, both those in the dispersion medium and on the particle surfaces, automatically rearrange themselves.

The density of ions on the particle surface can be regarded as one of the parameters which define the thermodynamic state of the system.* Let the number of adsorbed ions of species 1 per unit area of the inner surfaces of the two plates be denoted by ν_1, which we at first assume to be an arbitrary number, free to vary independently of λ. We shall ignore the influence of the outer surfaces of the two particles. Then the free energy associated with the ions of the whole system at stage λ has the form

---

[17] Fowler, *Statistical Mechanics*, 2nd. edition (Cambridge University Press, 1936).

[18] Onsager, *Chem. Rev*, 1933, 13, 73.

[19] Fuoss, *J. Chem. Physics*, 1934, 2, 818.

* Derjaguin[16] has referred to the densities of the charges in the ionic atmospheres of the colloidal particles as corresponding to ' uncontrolled ' co-ordinates, which are established on account of the thermal motion. The surface density of ions ν_1, however, is specified as controlled '.

$$F = F(\lambda, \nu_1, R\,;\; N_1, N_2, \ldots N_s, V, T), \qquad . \quad . \quad (2)$$

where N_i is the total number of ions of species i in the solution, there being s species in all, subject to the condition that

$$\frac{\partial F}{\partial \nu_1} = 0, \qquad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where all the other parameters, including R, are kept constant when differentiating with respect to ν_1. Equation (3) defines the average ionic density $\nu_1 = \overline{\nu_1}(\lambda, R)$ under equilibrium conditions. Henceforth it will be understood that the quantities $N_1, \ldots N_s$, V and T are kept constant and hence we need not explicitly write them down as parameters.

Onsager [18] has shown that it is still possible to apply the formula (1) when the configurations available to the ions are restricted in some way. This means that we evaluate the positional part of the Gibbs phase integral (assuming classical statistics) associated with the ions over a sub-space of the entire configurational phase-space. In particular, such a sub-space would be defined by specifying that ν_1 ions of type 1 be assigned to each unit area of the particle surfaces. The phase integral over this sub-space would yield the free energy expression $F(\lambda, \nu_1, R)$ at each stage λ in the charging process. Now, however, the averaging process in formula (1) will be subject to a fixed surface density of ions ν_1. Thus, in between the successive stages of charging, the ions in the dispersion medium would automatically re-arrange themselves into their new equilibrium state, but the number of adsorbed ions on the particle surfaces would remain unchanged.

To evaluate $F(\lambda, \nu_1, R)$ we first consider the discharged state $\lambda = 0$ with ν_1 discharged ions of type 1 on unit area of particle surface. It will be assumed that the free energy to be associated with this state, which we shall denote by $F_0(\nu_1)$, is a function of ν_1 (and also of $N_1, \ldots, N_s$, V and T, of course), but is independent of R. The usual treatment of adsorbed monolayers of molecules on surfaces [20] suggests that a good approximation to $F_0(\nu_1)$ can be written in the form *

$$F_0(\nu_1) = \sum_{i=1}^{s} N_i\chi_i^0 + 2\mathrm{A}\nu_1(-\,\mathrm{B} + kT\ln\nu_1 - \mathrm{C}\nu_1) \quad . \quad . \quad (4)$$

Here χ_i^0 is the chemical potential of a discharged i ion in the solution except if $i = 1$. χ_1^0 would be the chemical potential of a discharged ion of type 1 if all the N_1 ions were dissolved in the dispersion medium of volume V, so that none are adsorbed on the particle surface. χ_i^0, and the quantities B and C are functions of $N_i, \ldots, N_s$, V and T only, being independent of ν_1 and R. k is Boltzmann's constant. The main contribution to the term $-\,2\mathrm{AC}\nu_1^2$ comes from the interaction of pairs of discharged ions adsorbed on adjacent sites on the particle surfaces. We can expect this quantity to be unimportant if only a small fraction of the available sites are occupied by adsorbed ions. To simplify matters we shall neglect this term and make only brief reference to it near the end of this paper.

For given ν_1 and R it is possible to show, by using (1) that the increase in free energy of the whole system, associated with the charging process from o to λ, is

---

[20] See, for example, ref. 17 or Fowler and Guggenheim, *Statistical Thermodynamics* (Cambridge University Press, 1939).

* In the original presentation of this paper, the author had omitted the second term on the right-hand side of equation (4), which is the part that depends on ν_1. He was led to introduce this term after Dr. J. T. G. Overbeek, in a private communication, had correctly pointed out that the charging process was not being started in conditions that were independent of R.

$$F(\lambda, \nu_1, R) = \sum_{i=1}^{s} N_i \mu_i(\lambda, \nu_1, R) + 2A\nu_1(-B + kT \ln \nu_1)$$
$$- 2A\nu_1 \Delta\chi_1(\lambda) + 2A\int_0^{\lambda} \frac{\bar{E}(\lambda, \nu_1, R)}{\lambda} d\lambda. \quad (2i)$$

Here $$\mu_i(\lambda, \nu_1, R) = \chi_i^0 + \chi_i(\lambda) + e_i \int_0^{\lambda} \bar{\phi}_i(\lambda, \nu_1, R) d\lambda \;. \quad . \quad (5)$$

where $\chi_i(\lambda)$ is the electrical self-energy of an i ion, carrying the charge λe_i in the electrolyte solution and $\bar{\phi}_i(\lambda, \nu_1, R)$ is the average potential at an i ion due to the surrounding ions when the i ion is far removed from the two colloidal particles. This potential varies only very slightly with ν_1 and R, but nevertheless, this variation should be considered. We have introduced the notation

$$\Delta\chi_1(\lambda) = \chi_1(\lambda) - \chi_1^s(\lambda) \quad . \quad . \quad . \quad . \quad (6)$$

where $\chi_1^s(\lambda)$ is the electrical self-energy of an ion of species 1, carrying the charge λe_1, on the particle surface. $\mu_i(\lambda, \nu_1, R)$ is not quite the chemical potential of an i ion in the bulk of the solution owing to the particular manner in which χ_i^0 has been defined.

The quantity $\bar{E}(\lambda, \nu_1, R)$ is the average electrostatic potential energy of the ions in the two double layers associated with unit area of the inner sides of the parallel plates, at stage λ in the charging process. It is defined in the following way in (21). In the region of volume v between the two plates, we choose the x axis normal to the two parallel surface planes, the origin being at the median plane. Let $\bar{n}_i(\lambda, x)$ be the average volume density of ionic species i at position x and $\bar{\Psi}_i(\lambda, x)$ the average electrostatic potential at an i ion due to the surrounding ions. Both these quantities are functions of ν_1 and R, of course, but for the sake of simplicity in notation, we shall not explicitly indicate the dependence on these parameters. Instead of working with $\bar{\Psi}_i(\lambda, x)$, we shall introduce the change in potential $\bar{\psi}_i(\lambda, x)$, at an i ion, when it is brought from the bulk of the solution to position x, i.e.

$$\bar{\Psi}_i(\lambda, x) = \bar{\psi}_i(\lambda, x) + \bar{\phi}_i(\lambda, \nu_1, R) \quad . \quad . \quad . \quad (7)$$

The corresponding change in potential at an ion of kind 1 adsorbed on the inner surface of a plate will be $\bar{\psi}_1(\lambda, R/2)$. Then

$$\bar{E}(\lambda, \nu_1, R) = \lambda e_1 \nu_1 \bar{\psi}_1(\lambda, R/2) + \int_0^{R/2} dx \sum_{i=1}^{s} \bar{n}_i(\lambda, x) \lambda e_i \bar{\psi}_i(\lambda, x) \quad . \quad (8)$$

It is to be observed that ν_1 is assumed constant when integrating with respect to λ in equation (2i).

To obtain the dependence of $\mu_i(\lambda, \nu_1, R)$ on ν_1 and R we proceed as follows. Suppose that the electrolyte concentration in the solution is so dilute that we can use the first approximation to the potential function $\bar{\phi}_i(\lambda, \nu_1, R)$ in the Debye-Hückel theory of electrolytes, namely

$$\bar{\phi}_i(\lambda, \nu_1, R) = -\frac{\lambda e_i}{D} \kappa(\lambda, \nu_1, R), \quad . \quad . \quad . \quad (9)$$

where D is the dielectric constant of the solution and $\kappa(\lambda, \nu_1, R)$ is the characteristic quantity in the electrolyte theory referred to the parts of the solution distant from the colloidal particles. It is possible to express $\kappa(\lambda, \nu_1, R)$ in the form

$$\kappa(\lambda, \nu_1, R) = \kappa_0(\lambda) - \frac{4\pi\lambda^2}{DkTV}\frac{A}{\kappa_0(\lambda)}\left\{e_1^2\nu_1 + \sum_{i=1}^{s} e_i^2\int_0^{R/2} \overline{c}_i(\lambda, x)\mathrm{d}x\right\} \qquad (10)$$

very nearly, where

$$\kappa_0^2(\lambda) = \frac{4\pi\lambda^2}{DkTV}\sum_{i=1}^{s} N_i e_i^2.$$

$\overline{c}_1(\lambda, x)$ is the difference between the density of ions at position x and the density in the bulk of the solution. In obtaining (10), we retain terms proportional to the ratio v/V, but neglect the terms of higher order. If we now write

$$\overline{\phi}_i^0(\lambda) = -\frac{\lambda e_i}{D}\kappa_0(\lambda), \quad . \quad . \quad . \quad . \quad . \quad (11)$$

then the formula (2i) for the free energy can be expressed as:

$$F(\lambda, \nu_1, R) = \sum_{i=1}^{s} N_i\mu_i^0(\lambda) + 2\mathrm{A}\nu_1(-\mathrm{B} + kT\ln\nu_1)$$

$$- 2A\nu_1\Delta\chi_1^*(\lambda) + 2A\int_0^\lambda \frac{\overline{E^*}(\lambda, \nu_1, R)}{\lambda}\mathrm{d}\lambda, \quad (2ii)$$

where

$$\mu_i^0(\lambda) = \chi_i^0 + \chi_i(\lambda) + e_i\int_0^\lambda \overline{\phi_i}^0(\lambda)\mathrm{d}\lambda, \quad . \quad . \quad . \quad (5i)$$

$$\Delta\chi_1^*(\lambda) = \chi_1(\lambda) - \chi_1^s(\lambda) - \tfrac{1}{2}e_1\int_0^\lambda \overline{\phi_1}^0(\lambda)\mathrm{d}\lambda \quad . \quad . \quad (6i)$$

and

$$\overline{E^*}(\lambda, \nu_1, R) = \overline{E}(\lambda, \nu_1, R) + \overline{E'}(\lambda, \nu_1, R) \quad . \quad . \quad (8i)$$

where

$$\overline{E'}(\lambda, \nu_1, R) = \tfrac{1}{2}\frac{\lambda^2\kappa_0(\lambda)}{D}\sum_{i=1}^{s} e_i^2\int_0^{R/2}\overline{c_i}(\lambda, x)\mathrm{d}x. \quad . \quad . \quad (12)$$

It is to be observed that both $\mu_i^0(\lambda)$ and $\Delta\chi_1^*(\lambda)$ are functions of λ only, being independent of ν_1 and R. On the other hand, $\overline{c_i}(\lambda, x)$ does depend on ν_1 and R.

The equilibrium condition (3) now becomes

$$\Delta\chi_1^*(\lambda) + \mathrm{B} - kT - kT\ln\nu_1 = \frac{\partial}{\partial\nu_1}\int_0^\lambda \frac{\overline{E^*}(\lambda, \nu_1, R)}{\lambda}\mathrm{d}\lambda, \quad . \quad (3i)$$

which defines the function $\nu_1 = \overline{\nu_1}(\lambda, R)$. We need not consider the first term on the right-hand side of (2ii) since it is independent of R. Thus the remaining terms in (2ii), subject to the condition (3i) represent the interaction energy of the two plate-like particles. In integrating with respect to λ in (2ii) and (3i), it is understood that ν_1 is treated as a parameter independent of λ.

When $\lambda = 0$, the equilibrium condition (3i) defines the average surface density of discharged ions, say $\overline{\nu_1}(0)$, namely

$$kT\ln\overline{\nu_1}(0) = \mathrm{B} - kT \quad . \quad . \quad . \quad . \quad (3ii)$$

putting $\Delta\chi_1^*(0) = 0$. The free energy expression (2ii) reduces to

$$F_0(\overline{\nu_1}(0)) = \sum_{i=1}^{s} N_i\chi_i^0 - 2\mathrm{A}\overline{\nu_1}(0)kT \quad . \quad . \quad . \quad (4i)$$

Thus the interaction energy is independent of R at the start of the charging process.

Making use of (3i), the force between the two plates is

$$-\frac{\mathrm{d}F(\lambda, \overline{\nu_1}(\lambda, R), R)}{\mathrm{d}R} = -2\mathrm{A}\left[\frac{\partial}{\partial R}\int_0^\lambda \frac{\overline{E^*}(\lambda, \nu_1, R)}{\lambda}\mathrm{d}\lambda\right]_{\nu_1=\overline{\nu_1}(\lambda, R)} \quad . \quad (13)$$

A particularly simple result is obtained if we assume that $\overline{E^*}(\lambda, \nu_1, R)$ is proportional to ν_1^2, i.e. it can be expressed as

$$\overline{E^*}(\lambda, \nu_1, R) = \nu_1^2\overline{E_0^*}(\lambda, R). \quad . \quad . \quad . \quad (14)$$

Let us also introduce the potential function $\Phi_1^*(\lambda, R)$ defined by

$$\Delta\chi_1^*(\lambda) + \mathrm{B} - kT - kT \ln \overline{\nu_1}(\lambda, R) = \lambda e_1\Phi_1^*(\lambda, R) \quad . \quad . \quad (15)$$

and the average surface charge density

$$\bar{\sigma}(\lambda, R) = \lambda e_1\overline{\nu_1}(\lambda, R)$$

Then it is readily verified that the part of the interaction energy (2ii), which depends on R, may be expressed as

$$\begin{aligned} &F(\lambda, \overline{\nu_1}(\lambda, R), R) - F_0(\overline{\nu_1}(0)) \\ &= - A\overline{\nu_1}(\lambda, R)[\Delta\chi_1^*(\lambda) + \mathrm{B} + kT - kT \ln \overline{\nu_1}(\lambda,R)] + 2\mathrm{A}\overline{\nu_1}(0)kT \quad (16) \\ &= - A\bar{\sigma}(\lambda, R)\,\Phi_1^*(\lambda, R) - 2\mathrm{A}\,kT\,\{\overline{\nu_1}(\lambda, R) - \overline{\nu_1}(0)\}, \end{aligned}$$

where, according to (3i), the surface density function $\bar{\nu}_1(\lambda, R)$ (or $\bar{\sigma}(\lambda, R)$) is defined by :

$$\frac{\bar{\sigma}(\lambda, R)}{(\lambda e_1)^2} = \frac{\overline{\nu_1}(\lambda, R)}{\lambda e_1} = \frac{\Phi_1^*(\lambda, R)}{G\,(\lambda, R)} \quad . \quad . \quad . \quad (17)$$

where

$$G(\lambda, R) = 2\int_0^\lambda \frac{\overline{E_0^*}(\lambda, R)}{\lambda}\mathrm{d}\lambda \quad . \quad . \quad . \quad . \quad (17\mathrm{i})$$

The force between the two plates now becomes

$$- 2\mathrm{A}\bar{\nu}_1^2(\lambda, R)\frac{\mathrm{d}G(\lambda, R)}{\mathrm{d}R} \quad . \quad . \quad . \quad . \quad (13\mathrm{i})$$

The particular form (14) may be expected to hold for small values of ν_1, i.e. for small surface potentials. Under these circumstances, the results of Derjaguin,[16] Langmuir,[8] and Verwey [13] all agree, their common expression for the interaction free energy being $- \mathrm{A}\bar{\sigma}(\lambda, R)\,\Phi_1(\lambda)$ where $\Phi_1(\lambda)$ is the "surface potential". Our formula (16) for the energy differs from that of these authors in three respects. Firstly, our equation (17) which defines the charge density introduces distinctly new features. Secondly, their potential function $\Phi_1(\lambda)$ is replaced in our theory by $\Phi_1^*(\lambda, R)$, which depends on the separation R. Finally we have an additional term which is completely absent in their results. Our result is independent of any particular theory used to evaluate the potential functions $\bar{\psi}_\iota(\lambda, x)$. Furthermore, if we assume that $\Delta\chi_\iota^*(\lambda)$ and B are known, then our theory automatically yields the dependence of $\bar{\sigma}(\lambda, R)$ on λ and R. In other words, no special assumptions are required concerning the variation of surface potential with λ and R.

The dependence on ν_1 and R of the chemical potential $\mu_\iota(\lambda, \nu_1, R)$ of an ion distant from the particles is attributed to the variation of the electrolyte concentration in the bulk of the solution with interparticle distance R and surface charge density ν_1. This results in an extra contribution to the mutual free energy of the colloidal particles. For two particles immersed in a very large volume of electrolyte, one could assume that the electrolyte concentration at remote distances from the particles would not be affected by change in R or ν_1. However, in an actual sol, there will be a large number of particle pairs, forming Schiller layers or soap micelles for example. It then becomes obvious that the dependence of $\mu_\iota(\lambda, \nu_1, R)$ on R and ν_1 becomes more pronounced the more concentrated the sol. Indeed, it can be shown that our analysis can only apply to very dilute sols for we are actually expanding the chemical potential $\mu_\iota(\lambda, \nu_1, R)$ in powers of the sol concentration and retaining terms up to the first power in this concentration. Thus, in the case of soap solutions, where the concentration can be as much as 50 %, the higher terms in this expansion should be considered.

The relation between our analysis and that of Verwey can perhaps be understood in the following way. If, instead of keeping ν_1 fixed, we assume that ν_1 is given its equilibrium value $\overline{\nu_1}(\lambda, R)$ at each stage in the charging process, then the free energy expression (4) is replaced by

$$F(\lambda, \overline{\nu_1}(\lambda, R), R) = \sum_{i=1}^{s} N_i(\mu_i^{\circ}(\lambda) - \chi_i^{\circ})F_0(\overline{\nu_i}(0)) - 2A\int_0^{\lambda} \overline{\nu_1}(\lambda, R)\frac{\partial \Delta\chi_1^*(\lambda)}{\partial\lambda}d\lambda$$
$$+ 2A\int_0^{\lambda}\frac{\overline{E^*}(\lambda, \overline{\nu_1}(\lambda, R), R)}{\lambda}d\lambda. \quad (18)$$

This is essentially the form considered by Verwey who, however, neglected the variation of the chemical potential $\mu_i(\lambda, \nu_1, R)$ with respect to ν_1 and R. It can be shown that the formula (2ii), subject to the condition (3), is equivalent to (18). This is done by imagining the charging process from λ to $\lambda + d\lambda$ to be carried out in two steps : (i) the density of adsorbed ions $\overline{\nu_1}(\lambda, R)$ is held fixed while charging takes place, (ii) $\overline{\nu_1}(\lambda, R)$ is readjusted to its new equilibrium values at constant λ. The condition is now applied that the overall increase in free energy in step (ii) will be zero for all λ.

Following Verwey, let us neglect the variation of $\mu_i(\lambda, \nu_1, R)$ with R and ν_1 and assume the chemical potential of an i ion distant from the particles is $\mu_i^{0}(\lambda)$ defined by (5i). Then the condition that the chemical potential of ion of type 1 in the solution is equal to its chemical potential $\mu_1^{s}(\lambda)$, say, when adsorbed on the particle surface, may be written as :

$$\mu_1(\lambda, \nu_1, R) \approx \mu_1^{0}(\lambda) = \mu_1^{s}(\lambda) \quad . \quad . \quad . \quad (19)$$

valid for all values of λ. Now the replacement of $\mu_1(\lambda, \nu_1, R)$ by $\mu_1^{0}(\lambda)$ is equivalent to omitting the asterisk superscript wherever it appears in the formulæ given above. It will be understood that this has been done in the following discussion. Physically this means that we suppose the two plates are immersed in an electrolyte solution of infinite volume. Making use of (5i), the condition (19) may be written as

$$\mu_1^{s}(\lambda) = \chi_1^{0} + e_1\int_0^{\lambda} \bar{\phi}_1^{0}(\lambda)d\lambda + \chi_1^{s}(\lambda) + \Delta\chi_1(\lambda) \quad . \quad (19i)$$

Verwey proceeds to make two assumptions, namely

$$\Delta\chi_1(\lambda) = \lambda e_1\overline{\psi}_1(\lambda, R/2) = \lambda e_1\Phi_1(\lambda) \quad . \quad . \quad . \quad (20)$$

where his potential difference $\Phi_1(\lambda)$ between the solution and the particle surface is independent of R. The relation between $\Phi_1(\lambda)$ and our function $\Phi_1(\lambda), R)$, which is defined by (15) when the asterisk is omitted, is obtained from the equation

$$\Delta\chi_1(\lambda) = \frac{1}{2A}\left[\frac{dF_0(\nu_1)}{d\nu_1}\right]_{\nu_1=\overline{\nu_1}(\lambda, R)} + \lambda e_1\Phi_1(\lambda, R) \quad . \quad . \quad (21)$$

Now $\overline{\psi_1}(\lambda, R/2)$ is the mean potential at an adsorbed ion due to the neighbouring ions. In the expression (8) for the potential energy function $\overline{E}(\lambda, \nu_1, R)$ which we require in order to calculate the free energy, it is the mean potential that enters. On the other hand, the following argument suggests that $\Phi_1(\lambda, R)$ is the potential of the mean force acting on an adsorbed ion. Suppose that an ion of type 1 is brought from the interior of the solution to the particle surface, on which there are already $\overline{\nu_1}(\lambda, R)$ ions per unit area. We separate the " ideal " or non-electrical part of the free energy from the electrical part and imagine that a discharged ion is brought to the particle surface, carrying the density $\overline{\nu_1}(\lambda, R)$. The change in free energy to be attributed to this process is represented by the first term on the right-hand side of equation (21). The corresponding energy change resulting from the electrical forces would consist of two terms, the change in self-energy $-\Delta\chi_1(\lambda)$ and the work done against the mean Coulomb forces exerted on the ion by the surrounding ions. This latter work is defined by the potential of the mean force associated with the Coulomb interaction of the ions. The zero reference for this energy term is assumed to be the corresponding part of the chemical potential $\mu_1{}^0(\lambda)$ of an ion of type 1 in the solution, namely $e_1\int_0^\lambda \bar{\phi}_1{}^0(\lambda)d\lambda$. The equation (21) then states that the total free energy change in the process just described must be zero.

The potential of the mean force acting on an adsorbed ion, which we represent by $\Phi_1(\lambda, R)$, is not identical with the mean potential $\overline{\psi_1}(\lambda, R/2)$. Also, it appears that Verwey's function $\Phi_1(\lambda)$ is neither the mean potential nor the potential of the mean force. Our analysis indicates that Verwey did *not distinguish* carefully enough between the various *terms in* the chemical potential $\mu_1{}^s(\lambda)$ of an adsorbed ion. Thus, the first term on the right-hand side of equation (21), which represents a part of the non-electrical contribution to $\mu_1{}^s(\lambda)$, depends on the separation R through the term $2A\nu_1kT \ln \nu_1$ in $F_0(\nu_1)$. Physically, this quantity describes an entropy effect associated with the number of configurations that the ν_1 adsorbed ions can assume on unit area of particle surface. A similar entropy term is, of course, present in the Debye-Huckel theory of ordinary electrolytes, but is included in the ideal part of the free energy of the solution. This is also the case in the theory presented here. Since the equilibrium value of the surface density $\bar{\nu}_1(\lambda, R)$ depends on the distance R it follows that our entropy effect, which apparently was overlooked by Verwey, is now a function of R. It is the existence of this entropy term which leads to the dependence on R of the potential of the mean force.

We can explain why such an entropy term exists in another way. The work done against the interionic forces in bringing an ion to a given location on the particle surface is derived from the potential of the mean force. This work can be considered as the contribution from the Coulomb interactions to the " free energy " associated with an ion of type 1 being fixed at a definite position on the surface, the zero reference being the same ion in the interior of the solution. When we apply the condition (19), however, the transfer of an ion from the solution to the surface not merely involves placing this ion at a given site on the surface, but also permitting it to occupy any of the available sites, and consequently to interchange positions with the other adsorbed ions. This introduces an extra free energy accounting for the additional number of configurations available to all the adsorbed ions when their number is increased by 1. This extra free energy is a function of the density of the adsorbed ions, and hence of the interparticle distance.

Verwey also introduces the Poisson-Boltzmann equation in which the potential of the mean force is identified with the mean potential. Nevertheless, it is possible that by means of the above approximations, together with the use of the Poisson-Boltzmann equation, Verwey has been able to derive an expression for the mutual free energy which can explain many properties of colloidal solutions. However, when two colloidal particles are at some distance apart, the mutual interaction is obtained by subtracting two quantities which are very nearly equal, thereby obtaining a quantity of smaller order of magnitude. In doing this, the various assumptions discussed above may no longer be permissible and the dependence of both the mean potential $\bar{\psi}(\lambda, R/2)$ and the potential of the mean force $\Phi_1(\lambda, R)$ on the distance R, even if quite small, may play a decisive role.

3. Approximate Solution for Small Surface Potentials.

Suppose that the potential function $\overline{\psi}_i(\lambda, x)$ is the same, say $\overline{\psi}(\lambda, x)$, for the different types of ions. Assuming that $\overline{\psi}(\lambda, x)$ satisfies Poisson's equation :

$$\frac{d^2\overline{\psi}(\lambda, x)}{dx^2} = -\frac{4\pi}{D}\overline{\rho}(\lambda, x) = -\frac{4\pi}{D}\sum_{i=1}^{s} \lambda e_i \overline{n}_i(\lambda, x) \quad . \quad . \quad (22)$$

and making use of Green's theorem, (8) can be written in the standard forms

$$\overline{E}(\lambda, \nu_1, R) = \sigma\overline{\psi}(\lambda, R/2) + \int_0^{R/2} \overline{\psi}(\lambda, x)\overline{\rho}(\lambda, x)dx$$
$$\frac{D}{4\pi}\int_0^{R/2}\left[\frac{d\overline{\psi}(\lambda, x)}{dx}\right]^2 dx, \quad (23)$$

where the boundary condition to be satisfied by $\overline{\psi}(\lambda, x)$ at the particle surface is assumed to be

$$\sigma = \lambda e_1 \nu_1 = \frac{D}{4\pi}\left[\frac{d\overline{\psi}(\lambda, x)}{dx}\right]_{x=R/2}. \quad . \quad . \quad . \quad (24)$$

We shall not distinguish between $\kappa(\lambda, \nu_1, R)$ and $\kappa_0(\lambda)$ since they are almost identical and shall write $\kappa_0(\lambda) = \lambda\kappa_0$. It will finally be assumed that $\overline{\psi}(\lambda, x)$ satisfies the Poisson-Boltzmann equation, which for small enough values of $\overline{\psi}(\lambda, x)$ (25 millivolts) is approximated by

$$\frac{d^2\overline{\psi}(\lambda, x)}{dx^2} = \lambda^2\kappa_0^2\overline{\psi}(\lambda, x). \quad . \quad . \quad . \quad . \quad (25)$$

The nature of the approximations introduced above have been examined in some detail by Fowler,[17] Onsager,[18] and Kirkwood.[21]

The solution of (25) is

$$\overline{\psi}(\lambda, x) = \frac{4\pi e_1\nu_1}{D\kappa_0}\frac{\cosh \lambda\kappa_0 x}{\sinh \frac{1}{2}\lambda\kappa_0 R} = \overline{\psi}(\lambda, R/2)\frac{\cosh \lambda\kappa_0 x}{\cosh \frac{1}{2}\lambda\kappa_0 R}, \quad . \quad (26)$$

making use of (24). Substitution into (23) yields

$$\overline{E}(\lambda, \nu_1, R) = \frac{2\pi\lambda e_1^2}{D\kappa_0}\nu_1^2\left[\coth \tfrac{1}{2}\lambda\kappa_0 R - \frac{\lambda\kappa_0 R}{2\sinh^2 \frac{1}{2}\lambda\kappa_0 R}\right]. \quad . \quad (27)$$

If $\overline{n}_i^0$ is the average density of ions i far removed from the particles, we shall assume

$$\overline{c}_1(\lambda, x) = \overline{n}_i^0\{e^{-\lambda e_i\overline{\psi}(\lambda, x)/kT} - 1\}. \quad . \quad . \quad . \quad (28)$$

If we expand the exponential term in (28) and retain the first three terms in the expansion, then for a binary symmetrical electrolyte ($e_1 = e_2$), (13) becomes

$$\overline{E}'(\lambda, \nu_1, R) = \frac{\lambda^5\kappa_0}{4D(kT)^2}\left(\sum_{i=1}^{2}\overline{n}_i^0 e_i^4\right)\int_0^{R/2}\overline{\psi}^2(\lambda, x)dx$$
$$= \frac{2\pi e_1^2\lambda^4}{D\kappa_0}\left(\frac{e_1^2\kappa_0}{4DkT}\right)\nu_1^2\left[\coth \tfrac{1}{2}\lambda\kappa_0 R + \frac{\lambda\kappa_0 R}{2\sinh^2 \frac{1}{2}\lambda\kappa_0 R}\right]. \quad (29)$$

If we now add (27) and (29), an expression is obtained for $\overline{E^*}(\lambda, \nu_1, R)$ which is proportional to ν_1^2.

Now we are interested in the change in free energy as the parallel plates are brought closer together. Substituting (27) and (29) into (8i), and making use of the notation (14) and (17i), we obtain

$$G(\lambda, R) - G(\lambda, \infty)$$
$$= \frac{4\pi e_1^2}{D\kappa_0}\left[\coth \tfrac{1}{2}\lambda\kappa_0 R - \frac{2}{\lambda\kappa_0 R} - 1 + \frac{e_1^2\kappa_0}{4DkT}\left(\frac{2}{\lambda\kappa_0 R}\right)^4 J(\tfrac{1}{2}\lambda\kappa_0 R)\right], \quad (30)$$

---

[21] Kirkwood, *J. Chem. Physics*, 1934, 2, 767.

where
$$J(u) = \int_0^u v^3\left(\coth v + \frac{v}{\sinh^2 v} - 1\right)\mathrm{d}v. \qquad (31)$$
This quantity is negative for large R (proportional to $-1/R$) and positive for small R (proportional to $1/R$). There is a minimum which is given approximately by
$$R_{\text{min.}} = \left(\frac{2e_1^2}{DkT\kappa_0}\right)^{\frac{1}{2}} = \frac{6\cdot6}{\gamma^{\frac{1}{4}}} \text{ Angstroms}, \qquad (32)$$
when $\lambda = 1$, provided $\kappa_0 R_{\text{min.}}/2 < 1$. Here γ is the electrolyte concentration in mol./l. for a 1-1 type and $T = 18°$ c.

The contribution to the free energy from the electrostatic potential energy function $\overline{\overline{E}}(\lambda, \nu_1, R)$ is represented by the first part of (31). This term is negative for all values of R and hence gives an attractive force. The contribution to the free energy from the function $\overline{\overline{E}}'(\lambda, \nu_1, R)$ is represented by the second part of (31), which is positive for all R and hence is responsible for the repulsive force at small distances. Physically, the interpretation is that for given ν_1, when the parallel plates are brought closer together, there is a migration of ions from the bulk of the solution to the region between the two plates.

It follows from (15) and (17) that the variation of the surface density of ions with separation is given by
$$\frac{\mathrm{d}\overline{\nu_1}(\lambda, R)}{\mathrm{d}R} = -\frac{\overline{\nu_1}(\lambda, R)\dfrac{\mathrm{d}G(\lambda_1 R)}{\mathrm{d}R}}{G(\lambda, R) + kT/\overline{\nu_1}(\lambda, R)} \qquad (33)$$
Noting that $G(\lambda, R)$ and $\overline{\nu_1}(\lambda, R)$ are necessarily positive quantities, the derivative (33) is positive at large R, negative at small R, with a zero precisely at the minimum of $G(\lambda, R)$, as given by equation (32). This means that as the plates are brought together from infinite separation, the surface ionic density at first increases, reaches a maximum and then decreases. In contrast, Verwey's results show the surface density to be steadily diminishing as the plates approach one another. The force between the two plates, defined by the expression (13i) is attractive at large R and repulsive at small R. The force is zero, i.e. the interaction energy has a minimum at the position defined by equation (32).

Bergmann, Low-Beer and Zocher[1] found that the Schiller layers in the WO_3 sols were separated by distances which ranged from 4500 A. to 2100 A. as the electrolyte concentration (NaCl) varied from 4×10^{-5} to 64×10^{-5} mol./l. According to (32), the position of the minimum decreases from 83 to 42 A. for this change in electrolyte concentration. Thus the separations found experimentally are about 50 times those predicted by equation (32). Furthermore, when the gravity effect is taken into account, the discrepancy is even greater.

If we consider the 25 % emulsion of lipide extracts in 0·12 % N NaCl, examined by Palmer and Schmitt,[3] the experimental separation was 85 A., whereas the formula (31) yields the value 11 A. However, the concentration of lipide is very high (80 % of the water is in the water layers) and at such small distances other effects such as image forces * and van der Waals forces would begin to play an important part. These factors have been neglected here. Palmer and Schmitt found that the distance R decreased as the electrolyte concentration was increased. In contrast, in the case of the soap micelles, Harkins, Mattoon and Corrin[8] observed that the addition of KCl or NaCl actually increased the separation of the micelles, in direct contradiction to our formula (32). Thus our theory is not applicable either to the concentrated soap solutions or to the lipide extract emulsions.

---

* These have been considered by Wagner, *Physik. Z.*, 1924, **25**, 474 and Onsager and Samaras, *J. Chem. Physics*, 1934, **2**, 528, in their theory of the surface tension of strong electrolytes.

We proceed to compare our results with those obtained by Verwey, who used an expression for the free energy equivalent to our formula (18). If we insert his modification (21) into (9) and replace $\overline{\psi}_i(\lambda, x)$ by $\overline{\psi}(\lambda, x)$, then we need to consider the free energy expression at stage λ

$$- 2A\int_0^\lambda \overline{\nu_1}(\lambda, R)\frac{\partial}{\partial\lambda}(\lambda e_1\Phi_1(\lambda))\mathrm{d}\lambda + 2A\int_0^\lambda \overline{\nu_1}(\lambda, R)e_1\Phi_1(\lambda)\mathrm{d}\lambda + 2A\int_0^\lambda \frac{\mathrm{d}\lambda}{\lambda}\int_0^{R/2} \overline{\psi}(\lambda, x)\overline{\rho}(\lambda, x)\mathrm{d}x. \quad (34)$$

If we now introduce for $\overline{\psi}(\lambda, x)$ the formula (25) with $\psi(\lambda, R/2)$ replaced by $\Phi_1(\lambda)$ and use (24) to obtain $\nu_1 = \overline{\nu_1}(\lambda, R)$, then it can be shown that (34) simply becomes

$$- \mathrm{A}\overline{\sigma}(\lambda, R)\Phi_1(\lambda) = - \frac{DA}{4\pi}\lambda\kappa_0\Phi_1^2(\lambda) \tanh \tfrac{1}{2}\lambda\kappa_0 R \quad . \quad . \quad (35)$$

and the mutual energy is

$$\frac{DA}{4\pi}\lambda\kappa_0\Phi_1^2(\lambda)(1 - \tanh \tfrac{1}{2}\lambda\kappa_0 R). \quad . \quad . \quad . \quad (36)$$

This expression yields a repulsive force between the two plates for all separations. Furthermore, Verwey has in this way obtained a solution which is self-consistent. Putting $\lambda = 1$, (36) becomes the result already obtained by Derjaguin [16] and Langmuir [8] by different arguments.

Finally, Verwey's theory leads to reasonable agreement with the experimental results of Bergmann, Low-Beer and Zocher on the separations of the WO_3 particles which form Zocher layers. This follows when we equate the repulsive forces between the plates to the force of gravity.

We shall only briefly refer to the neglected term $- 2\mathrm{AC}\nu_1^2$ in (4). If this is introduced, then it is possible for the interaction of the two plates to show a repulsion at large distances provided $\mathrm{C} > G(1, \infty)$. However, preliminary calculations indicate that such a large value of C is rather unlikely.

The theory that we have developed in the previous section is also self-consistent and in the general formulation avoids the approximations in equations (20), as introduced by Verwey. However, when the Poisson-Boltzmann equation is used to obtain an expression for $\overline{E_0}^*(\lambda, R)$, we run into difficulties. We find that the replacement of $\overline{\psi}_i(\lambda, x)$ by a common potential $\overline{\psi}(\lambda, x)$ and the subsequent use of the Poisson-Boltzmann equation implies the general equivalence of (34) and (35) (for small potentials), the surface density $\overline{\sigma}(\lambda, R)$ being defined by the usual boundary condition (24).

In spite of the apparent consistency of Verwey's results, there is one conclusion suggested by our theory which is at variance with Verwey's general result concerning the repulsion between colloidal particles. If the potential energy expression $\overline{E_0}^*(\lambda, R) - \overline{E_0}^*(\lambda, \infty)$ is negative for large R, then we can expect the same to be true of the free energy. However, if we use the Poisson-Boltzmann equation, and therefore the solution (26) for the electrical mean potential, then for two parallel plates the interaction potential energy is negative, whereas Verwey's result for the interaction free energy is positive at large R.

If we neglect the variation, with respect to ν_1 and R, of the chemical potential $\mu_i(\lambda, \nu_1, R)$ of an ion in the solution, then $\overline{E_0}^*(\lambda, R)$ is replaced by $\overline{E}_0(\lambda, R)$, which is the electrical field energy of the double layers per unit area of the plates and per unit charge density on the plates. This is essentially the quantity considered by Dube and the author in the earlier papers. For two spherical particles [14] and also for two parallel cylindrical particles,[15] the expression $\overline{E}_0(\lambda, R) - \overline{E_0}(\lambda, \infty)$ is negative at large R and positive at small R. Preliminary calculations have also been made on tactoids consisting of parallel and equidistant rod-like particles forming a regular hexagonal lattice in cross-section. A first approximation to $\overline{E_0}(\lambda, R)$ yields a repulsion term for all R but it is quite possible that more accurate calculations will yield an attraction at large separations in view of the results for two cylindrical particles. These calculations suggest that the effect of the finite size of the plates in Schiller layers is to produce a repulsive contribution.* It is perhaps significant that for soap micelles, to which the model of infinite parallel plates is more applicable than to the Schiller layers, the separation distance is much smaller. It is true, of course, that in our calculation of $\overline{E_0}(\lambda, R)$ for the different kinds of particles, we have been inconsistent in utilising the approximate equation (25). This point requires further study.

Verwey distinguished between the " electrical " free energy and the " chemical " free energy,† which for small enough potentials are represented in this paper by the quantities

$$2\mathrm{A}\overline{\nu_1}^2(\lambda, R)G(\lambda, R) \text{ and } -2\mathrm{A}\overline{\nu_1}(\lambda, R)[\Delta\chi_1{}^*(\lambda) + \mathrm{B} - kT \ln \overline{\nu_1}(\lambda, R)]$$

respectively. He stressed that when his potential difference $\Phi_1(\lambda)$ is independent of R, then the chemical free energy determines the character of the forces between the particles. He based this conclusion on the fact that when the electrical double layer is formed, the change in chemical energy is always greater than that in electrical energy. Thus, for small potentials, our theory shows that

$$-\text{ Chemical energy} = 2 \text{ (Electrical energy)} + \mathrm{A}\overline{\nu_1}(\lambda, R)kT \quad . \quad (37)$$

However, the interparticle forces are determined not by the magnitudes of these quantities, but rather by their rate of variation with respect to the distance R.

An essential difference between our theory and that of Verwey is that our results indicate that it is the electrical free energy which is responsible for the type of forces between the particles. A crucial feature of our method is that the chemical energy can be described by a function $F_c^*(\lambda, \nu_1)$ which depends on λ and ν_1, whereas the electrical energy has the form $F_e^*(\lambda, \nu_1, R)$ depending on λ, ν_1 and R. These properties are sufficient to ensure that the force between the particles is given by equation (13), i.e. we differentiate $F_e^*(\lambda, \nu_1, R)$ partially with respect to R and then substitute $\nu_1 = \overline{\nu_1}(\lambda, R)$. The proof for this simple theorem is as follows. Suppose that the particles are at a specified distance R and that the density of adsorbed ions under equilibrium conditions is $\nu_1 = \overline{\nu_1}(\lambda, R)$. This means that the overall change in free energy is zero when we vary $\overline{\nu_1}(\lambda, R)$ by an amount $\delta\bar{\nu}_1$. This yields our equation (3) which we can now write as

$$\left[\frac{\partial F_c^*(\lambda, \nu_1)}{\partial \nu_1} + \frac{\delta F_e^*(\lambda, \nu_1, R)}{\partial \nu_1}\right]_{\nu_1=\overline{\nu_1}(\lambda, R)} = 0 \quad . \quad . \quad . \quad . \quad (38)$$

Suppose now that R is changed by an infinitesimal amount δR. To restore equilibrium $\overline{\nu_1}(\lambda, R)$ is varied by an amount $\delta\overline{\nu_1}$. The change in free energy will be

$$\left[\frac{\partial F_c^*(\lambda, \nu_1)}{\partial \nu_1} + \frac{\partial F_e^*(\lambda, \nu_1, R)}{\partial \nu_1}\right]_{\nu_1=\overline{\nu_1}(\lambda, R)} \delta\overline{\nu_1} + \left[\frac{\partial F_e^*(\lambda, \nu_1, R)}{\partial R}\right]_{\nu_1=\overline{\nu_1}(\lambda, R)} \delta R \quad . \quad (39)$$

On substituting (38) this reduces to the last term. We have thus proved a general theorem which is quite independent of any approximations which may be introduced to evaluate the free energy, namely, the force between the particles is obtained directly from the electrical free energy.

A simple physical interpretation can be given to the form $F_c^*(\lambda, \nu_1)$ for the chemical free energy. This energy is associated with the adsorption of ions on the particle surface. If the surface density of ions is unchanged, then the chemical energy is also unchanged. The electrical free energy accounts for the long-range interionic forces, and consequently the separation of the particles enters into this energy directly. On the other hand, one can say that the separation R is introduced into the chemical energy indirectly through the variation of the equilibrium value of the ionic surface density with R. If the distance is varied at constant surface density, the electrical free energy changes but the chemical free energy does not. If the electrical forces are absent, then the chemical energy would not depend on the distance R at all, except when the particles are very close together ($R < 10$ A., say).

---

* Further investigation, presented in Part II, suggests that this repulsion may be considerable.

† The use of the term " chemical " in this case is perhaps unfortunate, since there is included the electrical self-energy of the ions.

The striking divergence between our results and those of Verwey, Langmuir and Derjaguin certainly suggests that an analysis should be made of the errors produced by applying the Poisson-Boltzmann equation to the double layers of colloidal particles. The inconsistencies in this equation when applied to the region near a phase boundary have been discussed recently by Coolidge and Juda.[22] It is precisely at the surface of the particles that the Poisson-Boltzmann equation, in conjunction with the boundary condition (24), is subject to error. But for small surface potentials the mutual free energy of two colloidal particles is directly proportional to the surface charge density $\bar{\sigma}(\lambda, R)$, which may be sensitive to the form of the equation used to determine the potential functions in the neighbourhood of the particle surface.

In conclusion, the results so far obtained do allow for the existence of long range attractive forces between the double layers of colloidal particles, without the introduction of the van der Waals-London dispersion forces. Further work on this problem is in progress.

The general theory developed in Section II of this paper is valid, although the particular form for the free energy of the discharged state as a function of ν_1 may only be applicable to certain cases. In the opinion of the author it is an improvement on the theory of Verwey and Overbeek in that it avoids the assumption that the surface potential is independent of the separation. However, the application to parallel plates must be discarded in its present form because of the anomalous fact that the potential in the region between the two plates rises to infinity as the ions are discharged, assuming the surface density of ions is kept fixed. Hence definite evidence for the existence of attractive forces between the electrical double layers of two colloidal particles has not been obtained. It should be remarked that the calculations in this paper only apply to very dilute sols and that it is far from certain that attractive forces of an electrical energy cannot exist in concentrated sols under suitable conditions.

Summary.

A method of calculating the free energy of interaction of two charged colloidal particles is outlined, with particular application to two parallel plates. The theory resembles that developed by Verwey but differs in the manner in which account is taken of the thermodynamic equilibrium between the ions adsorbed on the surface of the particles and those in the solution. We introduce the condition that the mutual free energy must be a minimum with respect to the surface density of adsorbed ions.

Our results differ from those obtained by Verwey, Langmuir and Derjaguin in that evidence is obtained for the existence of a minimum in the mutual free energy of two charged colloidal particles, without the introduction of the attractive London dispersion forces. In the case of two parallel plates, however, the present tentative calculations place this minimum at much smaller spacings than are found with Schiller layers.

[22] Coolidge and Juda, *J. Amer. Chem. Soc.*, 1946, 68, 608.

It is suggested that among the reasons for the discrepancy are the use of the Poisson-Boltzmann equation, and the neglect of the image effects and of the finite size of the colloidal plates. A general theorem relating to the mutual free energy, which is independent of the Poisson-Boltzmann equation, is obtained and seems to be inconsistent with Verwey's general conclusions concerning the repulsive character of the forces between colloidal particles for all distances.

The author is indebted to Prof. E. A. Guggenheim for a number of extremely helpful discussions on the theoretical basis of the free energy expressions (4) and (9), to Drs. E. J. W. Verwey and J. T. G. Overbeek for discussion and correspondence with regard to the role of the chemical energy of adsorption, and to Prof. J. D. Bernal for criticisms and suggestions.

Nuffield Research Laboratory,
Birkbeck College,
London.

LONG DISTANCE FORCES ACTING BETWEEN COLLOIDAL PARTICLES.

BY E. J. W. VERWEY AND J. TH. G. OVERBEEK.

Received 14th August, 1946.

The theory of the long distance force acting between the particles in a colloidal solution, more especially in the case where the particles are surrounded by an electrical double layer, was considered shortly before the war by several authors. Their results were contradictory in the main points. We have made an extensive investigation of the same problem,[1, 2] and arrived at the conclusion that the theoretical considerations of the authors referred to above were all more or less unsatisfactory.

In section 1 of the present paper we shall give a brief critical discussion of this work. In sections 2 and 3 we shall then give a very brief summary of our own theoretical considerations and the results we obtained.

1. Work of Previous Authors.

The work of Corkill and Rosenhead [3] and the extensive theory of Levine and Dube [4] have this point in common that they both arrive at the conclusion that the interaction of the double layers of two parallel plates or two spherical particles leads to an attraction, at least at larger distances. From a theoretical point of view their treatments are closely analogous. Their considerations are, however, incorrect for two reasons :

(*a*) The main objection against their considerations of the case of two colloidal particles for which the particle charge is not altered by the interpenetration of the double layers is that they calculate the energy of the electrical field of the double layer system. In reality, however, the force between the particles due to the interaction of the double layers

---

[1] (*a*) Verwey, *Trans. Faraday Soc.*, 1940, **36**, 192 ; (*b*) *Chem. Weekblad*, 1942, **39**, 563 (in Dutch) ; (*c*) Contribution to a symposium held at Utrecht by the *Ned. Chem. Ver.*, July, 1944, (in Dutch) ; (*d*) *Philips Research Reports*, 1945, **1**, 33 (English translation of 1(*c*)).

[2] Verwey and Overbeek, *Theory of the Stability of Lyophobic Colloids* (Elsevier, Amsterdam, 1946).

[3] Corkill and Rosenhead, *Proc. Roy. Soc. A.*, 1939, **172**, 310.

[4] Levine, *ibid.*, **170**, 145, 165 ; Levine and Dube, *Trans. Faraday Soc.*, 1939, **35**, 1125, 1141 ; 1940, **36**, 215 ; *J. Physic. Chem.*, 1942, **46**, 239 ; *Phil. Mag.*, 1940, (7), **29**, 105.

is determined by the variation of the *free* energy of the system as a function of the distance. Accordingly these authors neglected the entropy associated with the thermal motion of the ions in the solution carrying the charge of the diffuse outer layer. We have been able to show that the two quantities (field energy and free energy) behave in an entirely different way as a function of the distance. The correct theory leads to a *repulsive* force between the particles for all particle distances.

(*b*) Usually—especially in lyophobic systems—the potential of the double layer will be independent of the distance between the particles. In this case the charge of the particles is altered by the interaction of the double layers (i.e. it decreases with decreasing distance between the particles). This possibility has also been considered by the authors mentioned above. In this case their erroneous results are mainly explained by the fact that they do not take into consideration the chemical changes associated with the variation of the charge. It can easily be shown, however, that the change in the chemical free energy is the predominant term if the distance between the particles is varied while keeping the surface potential constant. The correct theory again leads to a *repulsion* for all particle distances as a consequence of the interaction between the double layers.

The work of Langmuir [5] has an entirely different character; he also arrives at the result that the ionic spheres of the particles are responsible for attractive forces between the particles. He considers the colloid as a whole (charged particles + dissolved ions) and applies the Debye-Huckel electrolyte theory to such a system. The result mentioned above, however, is entirely explained by the errors caused by the use of the linear Debye-Hückel approximation.

Derjaguin [6] and also Langmuir in the last part of his paper, derives an equation for the force between two parallel plates covered by an electrical double layer. These force equations are fundamentally correct and can also be derived from free energy considerations. Derjaguin extends his theory to spherical particles. His theory of the stability of lyophobic colloids, however, is unsatisfactory for two reasons.

(*a*) In order to reduce the mathematical difficulties, he applies the double layer theory in its linear approximation, analogous to the Debye-Hückel theory of electrolytes. For colloid systems, however, especially if the particles are no longer small in comparison with the "thickness" of the ionic spheres, this approximation gives results which are quantitatively completely unsatisfactory.

(*b*) In Derjaguin's theory the flocculation of a sol is considered to be induced by an infinitely large attraction force acting only when the particles are in direct contact. In our theory we have been able to show that the long distance action of the London-van der Waals attractive forces is an essential factor in determining the stability of colloids.

2. Outline of the Theory Presented Here.

The effect of the interaction of the double layers surrounding the particles is most easily found by considering the free energy of the systems as a function of the distance between the particles. This free energy can in principle be obtained in two different ways, which lead to the same result. The first is to start from a generalised Lippmann equation, giving the free energy as a function of the surface charges and the surface potential. For the simple case of two parallel plates we have per cm.^{2}

$$F = -\int_0^{\psi_0} \sigma d\psi_0 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

---

[5] Langmuir, *J. Chem. Physics*, 1938, 6, 893.

[6] Derjaguin and Kussakow, *Acta Psychicochim.*, 1939, 10, 25, 153. Derjaguin, *ibid.*, 1939, 10, 333. *Trans. Faraday Soc.*, 1940, 36, 203.

in which ψ_0 is the electrical potential of the plate surface and σ the surface charge density. For the integration of this equation we must know σ as a function of ψ_0; as we did not wish to apply the double layer theory in its linear approximation, this is a rather complicated function containing, for instance, the distance of the plates and the concentrations and the valencies of the ions in the medium.

The second way shows some analogy with the derivation of the free energy of an ion in its ionic sphere as given in Debye and Huckel's theory of electrolytes. According to this method the free energy of the double layer system is found as the total amount of work associated with the gradual discharge of all ions of the system, including the ions responsible for the surface charge. If this discharging process is carried out under such conditions that the surface potential is kept constant the chemical work exactly cancels the electrical work for the surface charge. The final expression, therefore, contains only the charge of the diffuse part of the double layer. Again, for the case of two parallel plates we have (when we only consider one half of the system between $x = 0$ and $x = H/2$, H being the distance between the plates)

$$F = \int_0^1 \frac{d\lambda}{\lambda} \int_0^{H/2} \rho' \cdot \psi' dx \quad . \quad . \quad . \quad . \quad (2)$$

in which ρ' and ψ' are respectively the space charge and the electrical potential in the solution corresponding to an ionic charge λe (during the discharging process λ decreases continually from 1 to 0). Accordingly ρ' and ψ' are functions of x (the distance from the surface), λ, and H.

Using the differential equation (for sake of simplicity we only consider symmetrical electrolytes, ionic valency $= v$):

$$\Delta\psi = \frac{8\pi nve}{\epsilon} \sinh (ve\psi/kT) \quad . \quad . \quad . \quad (3)$$

which also forms the starting point of the Gouy-Chapman theory for a single double layer, equation (2) can be solved and evaluated numerically. Thus we obtain F as a function of the distance between the plates, with v (the valency of the ions), n (the ionic concentration) and ψ_0 as variable parameters. From this we directly obtain the repulsive potential V_R per cm.^{2} due to the double layer interaction

$$V_R = 2(F - F_\infty)$$

in which F_∞ is the free energy for $d = \infty$.

The theory shows that the most important part of the potential curve V_R is that for which $\frac{\kappa H}{2} < 1$, in which κ is the well-known Debye quantity.

$$\kappa = \left(\frac{8\pi n}{\epsilon kT}\right)^{\frac{1}{2}} . ve.$$

For the distances in this region the repulsive potential function approaches an exponential decay function, exp. $(-\kappa H)$. The theory can also be extended to spherical particles by an approximative method already suggested by Derjaguin, applicable in those cases for which $\kappa a > 5$, if a is the particle radius. This condition is generally fulfilled in colloid systems containing some electrolyte in the sol medium. The potential curves obtained in this way do not differ much from those valid for the case of two flat plates.

The problem of the interaction of two spherical particles with a radius of the same order of magnitude as the double layer extension, or even smaller, cannot be solved in such a way, with the theory for two flat layers as a starting point. In this case it is unavoidable, for mathematical reasons, to simplify equation (3), and to introduce the linear approximation of Debye and Hückel, or, at best, the second or third approximation as has also been done by Levine and Dube. The free energy is now most easily obtained by making use of equation (1). A less satisfactory point of this theory is the use of the linear approximation. Fortunately, however, when the particles are small in comparison to the " thickness " of the double layer, the error involved by this simplification is considerably smaller than for larger particles.

Having thus obtained a fairly complete picture of the repulsive potential resulting from the interaction of the double layers, based on the same principles as the Gouy-Chapman theory for a single double layer * (diffuse double layer), the theory of the stability of colloids and the long distance forces acting between the particles can be completed by considering also the van der Waals-London attractive potential V_A.

As a first approximation we used here the well-known London equation stating that the attractive potential between two atoms is proportional to r^{-6}, in which r is the distance between the atoms. A fundamental point of London's theory is that the attraction between all atoms is additive. The summation can be substituted by an integration. This integration shows that the attractive potential between two flat plates decreases according to H^{-2} and therefore considerably less steeply than between two atoms (de Boer).[7] The attractive potential between two spheres is a rather complicated function (Hamaker); [8] for small distances it approaches a H^{-1} law, if H now means the closest distance between the surfaces of the spheres; for very large distances it will again be proportional to R^{-6}, if R is the distance between the centres, or approximately to H^{-6}.

The force acting between two colloidal particles can then be read from a diagram containing the total potential curve

$$V = V_R + V_A$$

in which V_R and V_A are both functions of the distance between the particles; V_A is a negative quantity.

London's theory of the dispersion forces is based on the conception that the attraction is due to the polarisation of one atom as a consequence of the fluctuating dipole in a second atom. It does not take into account the finite time needed for the transmission of the electromagnetic field from the first atom to the second, and vice versa. This retardation will come into effect if the distance between the particles becomes of the order of the characteristic wavelengths of the participating atoms, i.e. of the order of 1000 A. The van der Waals-London attractive potential between two colloidal particles some distance apart will generally contain many contributions of atomic pairs at such a distance or even at a larger distance. It can, therefore, be expected that the attractive potential especially at larger particle distances, will be appreciably smaller than suggested by the simple London-theory (Overbeek). Casimir and Polder have therefore reinvestigated the theory of dispersion forces (not yet published). They accounted for the effect mentioned above by applying the relativistic correction to the quantum mechanical theory. They found that for larger atomic distances the attractive potential between two atoms decays more rapidly than given by London's equation; the deviations become important for distances of the order of 500 A. and their results show that for large distances the attractive potential approaches a r^{-7} law instead of a r^{-6} law.

This correction has not yet been applied in the potential curves to be discussed in the following section. It should be borne in mind, however, that it may have a perceptible influence upon these curves especially for larger particle distances.

---

* We have also extended our theory for the interaction of two double layers of the Stern-type, in which a correction has been applied for the finite dimension of the ions in the medium. For this extension we must refer to the book mentioned in ref. 2.

[7] de Boer, *Trans. Faraday Soc.*, 1936, **32**, 21.

[8] Hamaker, *Physica*, 1937, **4**, 1058.

3. Results.

The total potential curves obtained according to the theory outlined in section 2 generally have the shapes indicated in Fig. 1 and 2. Fig. 1 gives a set of potential curves for different values of ψ_0 for two parallel plates in a rather concentrated solution of a monovalent electrolyte. Fig. 2 gives similar curves for two spherical particles in a more diluted electrolyte solution. The constant A is a proportionality factor determining the magnitude of the attractive potential; according to the quantum mechanical theory it will depend somewhat on the material of the particles and the sol medium but should be of the order of 10^{-12} erg. It is seen that

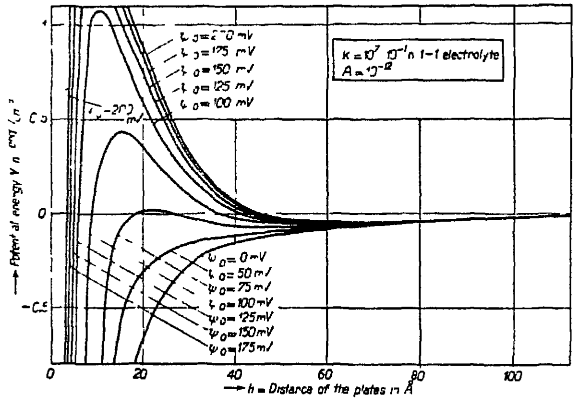


FIG. 1.

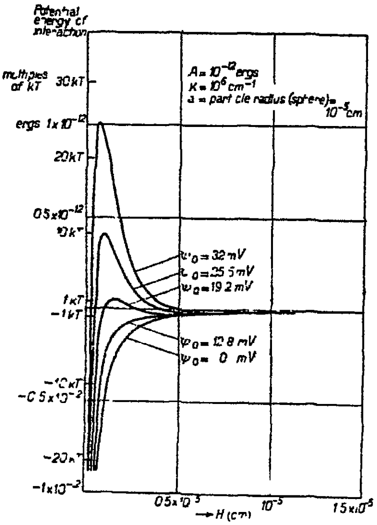


FIG. 2.

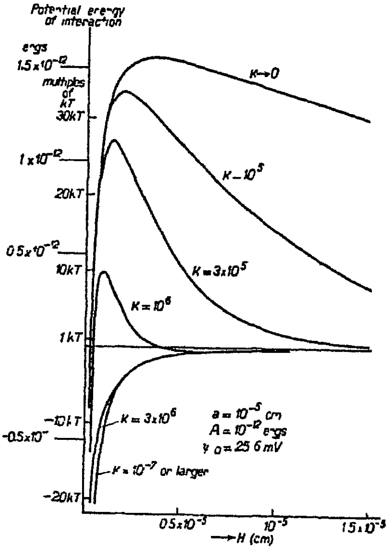


FIG. 3.

the double layer potential must have a certain minimum value in order to obtain a potential curve with a maximum; in the more dilute solution the potential required is lower than in the more concentrated electrolyte. This can also be read from Fig. 3, showing potential curves for $a = 10^{-5}$ cm. and $\psi_0 = 25 \cdot 6$ mv. for varying values of κ (κ is proportional to $\sqrt{n}$).

The theory for the kinetics of slow coagulation shows that in stable sols the potential barrier preventing the agglomeration must be at least 20 to 25 times kT. Hence in such a sol the particles cannot approach each other completely, and each particle is surrounded by a spherical liquid shell where a second particle of a similar kind never penetrates. The distance between the particle surfaces at which the potential curve has a maximum depends on the concentration and the valency of the ions in the sol medium, and also on the dimensions of the particles. If the particles are large in comparison with the " thickness " of the double layer the maximum is reached for a particle distance equal to about 1 to 2 times the characteristic length $1/\kappa$ (see Fig. 1 and 2) ; in " salt-free " water $1/\kappa$ is of the order of 1000 A., in 10^{-1} N solution 10 A. If the particles are smaller the maximum is reached for a smaller distance between the particle surfaces (see Fig. 3, curve for $\kappa = 10^6$). The presence of such a protective layer caused by the interaction of the double layers, which determines the stability of lypohobic colloids, is apparent from many other phenomena in colloid chemistry as e g. in the so-called electroviscous effect of lyophobic sols,* in the special arrangement in sediments from sols containing disc-shaped particles, in the phenomena observed in the formation of a deposit by electrophoresis, etc.

If the height of the maximum is decreased, which can be effected by decreasing ψ_0 (Fig. 1 and 2), or by increasing the electrolyte concentration (Fig. 3) the probability that a second particle crosses the potential barrier increases rapidly and the sol will show the phenomenon of a more or less " slow coagulation." For very low values of the maximum or for potential curves containing no maximum at all we finally reach the region of " rapid coagulation " in which every particle encounter leads to an agglomeration. In both cases the second particle enters the deep minimum for very short particle distances, i.e. the particles approach each other until the Born repulsive forces between the atoms of different particles come into action.

The theory shows that for short particle distances the van der Waals-London attraction potential generally prevails, causing the deep minimum which is responsible for the sticking together once the particles have come together within these small distances. All curves show however a second distance at which attraction always prevails. This is caused by the fact that the attractive potential, though decreasing more rapidly than the repulsive potential at short distances, decays more slowly at larger distances of the particles. Accordingly all curves containing a maximum at intermediate distances show a weak minimum at larger distances. This conclusion remains true if we apply the relativistic correction to the theory of dispersion forces, though the effect of this correction will be to lower the actual depth of these minima to an even lower level than suggested by the potential curves represented here. But even the minima occurring in Fig. 2 and 3 are so weak (the depth being less than kT) that they will normally escape experimental detection. Under very favourable conditions however, this minimum for large distances may come into effect and offer an explanation of certain phenomena observed in colloid chemistry. This may occur, for instance, for blade-shaped or linear particles oriented parallel to each other, as may be shown by comparing disc-shaped particles and spherical particles of the same radius. If we again take $\kappa = 10^6$, $a = 10^{-5}$ cm. and $\psi_0 = 25{\cdot}6$ mV. we find for two spherical particles a minimum of about 2×10^{-14} erg (about $\frac{1}{2}kT$), and for two disc-shaped particles oriented parallel and facing each other 2×10^{-13} erg, or about 5 kT (calculated from the equations for two flat plates). This minimum, corresponding to a distance between the particles of about 1000 A., has now sufficient depth to ensure that two particles with parallel orientation and at this distance from each other will travel together for a certain amount of time. There will also be a certain probability that this particle pair orients other particles in such a way as to form larger rows or groups in which the particles are captured in the field of two or more other particles. The depth of the minimum is proportional to a^2, and will therefore be larger for larger discs.

---

* The strong decrease of the viscosity of many lyophilic colloids caused by small amounts of electrolytes, though determined by the same repulsive forces between different parts of the same " particle " (molecule), is slightly more complicated due to a variation of the shape of the particles (contraction to a less stretched form).

A phenomenon as described in the foregoing paragraph has actually been observed in sols of Fe_2O_3 and V_2O_5 (a sol with lath-shaped particles) by Zocher.[9] This phenomenon which has been called the formation of tactoids is especially observed in aged sols, i.e. sols with sufficiently large particles. Similar phenomena have been described by Bernal and Fankuchen [10] for solutions of the tobacco mosaic virus, the particles of which have the shape of thin cylinders. Here the particles are arranged according to a two-dimensional hexagonal pattern. A parallel orientation, in such a way that the particles are in their weak potential minima, is in this case also favoured by the fact that each particle is surrounded by six direct neighbours. In order to test this hypothesis we have extended our theory to the interaction of cylindrical particles. A provisional calculation has shown that for a particle distance of 500 A. and a particle length of 5000 A. the minimum is at best of the order of magnitude of kT. As a somewhat deeper minimum seems to be required to explain the phenomena this might be an indication that the actual kinetic units are several times longer.

Natuurkundig Laboratorium der
N. V. Philips' Gloeilampenfabrieken,
Eindhoven,
Nederland.

---

[9] Zocher, *Z. anorg. Chem.*, 1925, **147**, 91.

[10] Bernal and Fankuchen, *Nature*, 1937, **139**, 923. *J. gen. Physiol.*, 1941, **25**, 111.

GENERAL DISCUSSION *

Dr. M. F. Perutz (*Cambridge*) said: I should like to draw attention to the different types of systems in which long-range forces appear to operate. With the exception of tobacco mosaic virus and V_2O_5, all structures so far investigated are of the laminar type. They form either continuous gels (montmorillonite), or expanded micelles (soaps), or tubular structures (phospholipoids). In the absence of any direct contact between the solute particles, the formation of regular periodic structures cannot be explained by the ordinary primary and secondary valency forces, except in certain special cases mentioned below. It is possible, at this stage, to distinguish three principal ranges (Table I).

TABLE I.

Range.		Examples.
1	< 20 A.	Certain types of soaps; *intra*-micellar swelling of clay minerals; crystalline proteins
2	20-150 A.	Soaps and phospholipoids
3	> 150 A.	Tobacco mosaic virus; *inter*-micellar swelling of clay minerals

Prof. Bernal has suggested that the first group is comparable to the so-called "kryohydrates," because the liquid, in some cases, appears to be frozen into a definite layer structure. Thus the liquid forms an ordered part of a crystalline or semi-crystalline system [1] whose regularity can be explained without invoking the presence of long-range forces. In protein crystals, moreover, the distance between the layers is independent of the concentration of neutral electrolyte, a fact which is incompatible with any of the theories of long-range forces so far put forward.

---

* On two preceding papers.

The greatest variety of phenomena has been discovered among the second group, and I should like to hear from the authors of the two previous papers to what extent their theories could be applied to those systems. Emulsions of mixed nerve lipoid, for instance, consist of tubular structures [2] in which the spacing between successive bi-molecular lipoid leaflets expands indefinitely, and nearly all available water is incorporated between those leaflets so long as salt is absent. In the presence of 0·12 N. NaCl, on the other hand, the liquid layers do not swell beyond a thickness of about 70 A., however much water is added.[3] Though this general type of behaviour is in accord with the predictions of Verwey and Overbeek's theory, the thickness of the liquid layer is much greater than their calculated value for flat plates at a similar salt concentration.

Colloidal soap solutions seem to contain disc-shaped micelles in which the distance between successive bi-molecular leaflets is inversely proportional to the soap concentration.[4] Surprisingly, addition of monovalent salts increases that distance.[5] The solubilisation of non-polar liquids in soap micelles is difficult to explain by any theory so far put forward. It might be suggested that purely osmotic forces drive the non-polar liquids into the micelles and the combination of surface tension and long-range London-van der Waals forces prevent the disintegration of the expanded micelles. Yet it is hard to see why such laminar structures should possess greater stability than globular emulsions. It is important to realise that neither soaps nor phospholipoids form expanded structures unless the hydrocarbon chains in the bi-molecular leaflets are in the smectic liquid-crystalline state.

The third range has been studied most extensively in gels of tobacco mosaic virus and montmorillonite. The particle distances may reach several thousand Angstroms and vary as an inverse function of the concentration of neutral electrolyte in the solvent.

With the application of more refined techniques it is to be expected that many instances of such regular structures will be discovered in biological systems. So far, no one has studied any of the systems in the second group with the express purpose of gaining information on the nature of the forces involved; it is probably difficult, therefore, to arrive at any general explanation, although the range from 20 to 150 A. may be biologically the most important.

Dr. F. Eirich (*Cambridge*) said: I would like to refer to a paper by F. London which, I feel, has not received sufficient attention, perhaps because the full paper appeared in a not very accessible periodical.[6] I had the opportunity to discuss the subject recently with Prof. London, and I think some of the emerging points may be of interest.

In his paper London states that in molecules containing extended electronic oscillators the spatial extension of the oscillators must be considered in addition to their dipole moment. Long oscillators, as in a system with conjugated double bonds, can be represented by monopoles of opposite sign suitably located in the molecule. The monopole forces may act at long range compared to the range usually associated with van der Waals forces.

---

[1] MacEwan, *Trans. Faraday Soc.*, this Discussion.

[2] Dervichian, *ibid.*

[3] Palmer and Schmitt, *J. Cell. Comp. Physiol.*, 1941, 17, 305.

[4] Hughes, Sawyer and Vinograd, *J. Chem. Physics*, 1945, 13, 131.

[5] Harkins, Mattoon and Corrin, *J. Amer. Chem. Soc.*, 1946, 68, 220.

[6] *Proc. 50th Anniv. Sympos. Chicago Amer. Assoc. Adv. Sci.*, 1941.*

* I thank Dr. Levine for telling me of a shorter version, since published in *J. Physic. Chem.*, 1942, 46, 305.

London's treatment differs from Overbeek and Verwey's in its approach, and also by being more specific. At first one would be inclined to think that there are practically no important colloidal systems to which London's results apply. However, London is inclined to think that charge transfer spectra and monopole forces are not necessarily confined to conjugated systems but may also obtain in " diluted " double bonds as in rubber, or localised ones as in polystyrene, and be responsible for a certain measure of behaviour normally attributed to overlapping Brillouin zones. If that be true, then I think extended oscillators may be expected in semi-conductors like tungstates or in transition element oxides of polymolecular crystal structures, or even generally on colloidal surfaces with ionogenic complexes whose structure permits resonance between ionic and covalent electron distributions. It would be of greatest interest to review these possibilities in the light of the recent work on Rochelle salt and perovskite anomalies.

Whether van der Waals attraction plays the decisive part in the formation of free energy minima, or whether Levine is right, must be left to a further critical study of theory and facts. If either is right I am, however, puzzled why the formation of structures is not more generally observed. It may be that the effects of long range forces have so far not been properly looked for and that we observed only the most conspicuous cases, while lesser ones like diffusion anomalies have escaped notice or the right interpretation. There remains also the problem of apparently long range forces and orders in the absence of Coulomb forces, as e.g. reported in " solubilisation." If such cases can be substantiated they will have to be reviewed, on account of the absence of long range repulsive van der Waals forces, in the light of osmotic phenomena, with long range attraction fulfilling the function of the membrane.

Mr. V. R. Gray (*Manchester*) said: The discussion on long distance forces between colloidal particles has so far been confined to those which can only operate in aqueous solutions, or in those possessing very high dielectric constants. There is thus no attempt to explain long distance forces operating in media of low dielectric constant.

An example of these forces is to be found in solutions of aluminium soaps in benzene and other hydrocarbons. For example, a solution of aluminium laurate in benzene when placed in an apparatus for measuring streaming birefringence of the coaxial cylinder type, displays birefringence when the system is at rest. The optic axis is parallel to the walls of the cylinder and the effect is not time dependent or affected by any previous movement of the apparatus. It is still capable of detection in quite dilute solutions, 0·2 % w./v. of soap. The amount of double refraction is less in the centre of the annulus, and decreases with rise of temperature. There must be a building up of layers of the soap chains from the walls which impose an orientation of neighbouring chains, continuing across the 2 mm. annulus.

This type of structure in aluminium soap suspensions in oils was noted by Derjaguin *et al.*[7] who, investigating the " wall-adjacent viscosity " of various oils, found that for oils thickened with aluminium naphthenate, " surface gel " formation occurred, that is, forming a layer close to the surface which behaved like a solid. It was however, only about 0·1 μ thick. It is evident that the forces acting in these systems can operate over quite considerable distances.

Dr. D. D. Eley (*Bristol*) said: I would suggest that the water structure in the neighbourhood of a protein molecule is altered over a shell several molecules thick, and that when two such shells overlap the total energy of the system changes, corresponding to a force between the protein molecules. This would appear to be a repulsive force, at first sight, but one cannot be certain as to its sign and magnitude without a theoretical investigation. An experimental investigation of water structure in protein gels and solutions would seem well worth while, from this point of view.

---

[7] Derjaguin, Koussakov and Krim, *Acta Physicochim.*, 1945, **20**, 35; Derjaguin, Strachovski and Malischeva, *ibid.*, 1944, **19**, 541.

Evidence in support of this suggestion derives from work on simple ionic solutions. Stewart[8] has considered X-ray patterns to show structural changes in the neighbourhood of the ion, probably extending over several molecular shells, and has interpreted partial ionic volume data on this basis. The heat capacity values of ions at infinite dilution would also seem to require long-range alterations in water structure.[9] Such effects of solvent structure will be in addition to Debye-Huckel ion-atmosphere effects.

Dr. J. Th. G. Overbeek (*Eindhoven*) (*partly communicated*): Dr. Levine offers an interesting method for calculating the interaction between two double layers. His results are fundamentally different from those obtained by Dr. Verwey and myself, in that he finds an attraction as the consequence of the interaction (at least for large separations) whereas we found a repulsion for any distance between the plates or particles. As the interaction of double layers seems to be one of the important causes for long range forces between small particles and large molecules it is thought important to analyses the reasons for this contradiction in some detail.

1. Dr. Levine includes in his analysis the influence of the finite dimensions of the whole system (*E'* in eq. (12)). As *E'* results in an extra repulsion the inclusion of those border-effects cannot explain the discrepancy.

2. In the end of section 2 of his paper Dr. Levine points to uncertainties involved in the use of the Poisson-Boltzmann equation and in the identification of the mean potential and the potential of the mean force. Kirkwood,[10] however, showed that the Poisson-Boltzmann equation may be used with confidence when the mean potential is small, and Casimir[11] could show in extending the analysis of Kirkwood that in the case of colloidal particles with many charges on their surfaces the Poisson-Boltzmann equation may even be applied for large potentials in the double layer.

Even the inconsistency of the Poisson-Boltzmann equation near a boundary surface suggested by Coolidge and Juda cannot explain the discrepancy, as the deviating result of Dr. Levine has been derived without applying any correction for the Coolidge-Juda effect.

3. In the end of section 3 it is argued that a negative value of $\bar{E}_0(R) - \bar{E}_0(\infty)$ would also imply a negative value of the change in *free* energy and consequently an attraction between the particles. This implication, however, is incorrect as it does not consider the influence of osmotic forces. We agree that for flat plates as well as for spherical particles the potential energy $\bar{E}$ is diminished by the interaction of the double layers. The increase in the ionic concentration in the double layer, however, which accompanies the interaction is the cause of a repulsion, and this repulsion always outweighs the attraction caused by the decrease of $\bar{E}$. This is not only proved by the consideration on the free energy as given by Verwey and myself but it follows also quite clearly from the analysis of the forces between the double layer as given by Derjaguin[12] and Langmuir.[13]

---

[8] Stewart, *J. Chem. Physics*, 1939, **7**, 869.

[9] Eley, *Trans. Faraday Soc.*, 1944, **40**, 184.

[10] Kirkwood, *J. Chem. Physics*, 1934, **2**, 767.

[11] Casimir, *Tweede symposium over sterke electrolyten en over de electrische ellaag*, 1944, p. 6, edited by the Ned. Chem. Ver., The Hague.

[12] Derjaguin, *Trans. Faraday Soc.*, 1940, **36**, 203.

[13] Langmuir, *J. Chem. Physics*, 1938, **6**, 873.

The only way out of the difficulty seems to be the presence of a mistake in Dr. Levine's derivations. I believe to have found such a mistake in his work, for if a simple correction is applied to equation (2i) the result obtained on the lines followed by Dr. Levine agrees exactly with our result.

In expressing the free energy of the Coulomb interaction Dr. Levine used the term $2A\int_0^\lambda \frac{\overline{E}\mathrm{d}\lambda}{\lambda}$, which describes the increase in Coulomb interaction during the charging process. It is, however, not certain that the free energy of the Coulomb interaction is zero when the charge of the ions approaches zero ($\lambda \rightarrow 0$). Although at first sight this seems a rather paradoxical statement, we meet such a case in the double layer of two parallel plates because the surface potential on the plates rises to infinity when the ions are discharged. (This follows immediately from equation (2b) relating the number of ions on the surface and the surface potential with the parameter, λ.)

The free energy of the Coulomb interaction for $\lambda \rightarrow 0$ cannot be derived by the method described by Dr. Levine, as in this method the number of ions on the surface (ν_1) is essentially kept constant.

It is, however, easy to derive a value for this free energy in considering another process, which starts from a surface without adsorption, and in which the (discharged) ions are moved one by one from the bulk of the solution to the surface, until a surface density ν_1 has been reached.

The work done in this process is equal to

$$\operatorname*{Lim}_{\lambda\rightarrow 0} 2A\int_0^{\sigma_0} \psi \mathrm{d}\sigma,$$

where ψ is the potential of the surface, σ its charge density and $\sigma_0 = \lambda e_1 \nu_1$, the density of charge finally reached.

Consequently, I propose to change equation (2i) into

$$\mathrm{F}(\lambda, \nu_1, R) = \sum_{i=2}^{s} N/\mu - 2A\nu_1\Delta\chi_1(\lambda) + 2A\int_0^\lambda \frac{\overline{E}(\lambda, \nu_1, R)}{\lambda}\mathrm{d}\lambda + \operatorname*{Lim}_{\lambda\rightarrow 0} 2A\int_0^{\sigma_0} \psi\mathrm{d}\sigma. \tag{2i corr.}$$

If now the influence of the finite dimensions of the system, being unimportant for the main problem is neglected, application of the equilibrium condition (3) to (2*i* corr.) leads to

$$\Delta\chi_1(\lambda) = \frac{\partial}{\partial\nu_1}\Big[\int_0^\lambda \frac{\overline{E}(\lambda, \nu_1, R)}{\lambda}\mathrm{d}\lambda + \operatorname*{Lim}_{\lambda\rightarrow 0}\Big\{\int_0^{\sigma_0} \psi\mathrm{d}\sigma\Big\}\Big]$$

which in the case of two flat plates is transformed into

$$\Delta\chi_1(\lambda) = \frac{\partial}{\partial\nu_1}\Big[\frac{4\pi e_1^2\nu_1^2}{D\kappa^2R}\{1/2\ \lambda\kappa R \coth 1/2\ \lambda\kappa R - 1\} + \frac{4\pi e_1^2\nu_1^2}{D\kappa^2R}\Big]$$

$$= \frac{4\pi e_1^2\nu_1}{D\kappa}\lambda \coth 1/2\ \lambda\kappa R = \lambda e_1\overline{\psi}(\lambda, R/2).$$

This is just the expression that has been refuted by Dr. Levine in his discussion on equation (20).

When the above expression for $\Delta\chi_1(\lambda)$ is introduced into (2*i* corr.) the free energy is found to be

$$\mathrm{F}(\lambda, \nu_1, R) = -DA/4\pi\ \lambda\kappa\psi^2(\lambda, R/2) \tanh \tfrac{1}{2}\lambda\kappa R$$

leading to a repulsion between the plates for any distance R and exactly agreeing with the results obtained by Verwey and myself.

In conclusion we may state that the method of Dr. Levine if applied correctly leads to the same results as our methods and affirms our conclusions that the interaction of electrical double layers always leads to a repulsion.

In reply to Dr. Perutz who asks to what extent our theory can be applied to systems of his second group, the minima in the potential curves, which are responsible for the formation of regular layer structure, are in our theory generally found at a separation between the layers of a few (i.e. up to 10) times $1/\kappa$, consequently distances may be expected ranging from a few tens of Angstroms (in solutions of 1/10 N or 1 N) to about 10 A. in pure water. The swelling of phospholipoid layers to a thickness of about 70 A. in 0·12 N NaCl is in good accord with the values of 50-70 A. that can be read from our Fig. 1. Moreover, the variation of the thickness of the liquid layer with the concentration of salts[14] follows easily from our theory. So in my opinion the behaviour of phospholipoids should be explained according to our theory. The same is the case with the intermicellar swelling of clay and with the formation of " bottom-layers " in solutions of tobacco mosaic virus.

The case of soap micelles, however, cannot be explained by our theory in its present state. In the soap solutions the concentration of free electrolyte (free soap ions) is probably very low, so that we should expect the water layers in the micelles to be much thicker than have been found. Moreover, as Dr. Perutz mentioned, the influence of the addition of salts on the spacing is in the wrong direction.

Dr. S. Levine (*London*) (*communicated*) : Dr. Overbeek correctly points out that, we meet with a paradoxical situation in the case of two parallel plates because the surface potential on the plates rises to infinity when the ions are discharged, assuming that the density of adsorbed ions is kept fixed. I cannot agree, however, with his solution to this difficulty. The notation used below will be the same as in my paper.

Dr. Overbeek states that there is a finite free energy of the Coulomb interaction when $\lambda \rightarrow 0$. Now this free energy is obtained by charging all the ions at the same rate, including those responsible for the charge on the plates. When $\lambda \rightarrow 0$ and the surface density of ions ν_1 is held constant, then the solution of the Poisson-Boltzmann equation yields a potential which becomes equal throughout the region between the two plates and which tends to infinity. Since the charge in the diffuse layers is equal and opposite to that on the plates, this ensures that our formal expression for this free energy tends to zero as $\lambda \rightarrow 0$. Nevertheless, the infinite value of this common potential in the case of two plates does create a difficulty which does not exist in other cases.

The implicit assumption is made in my theory that when the system is discharged, the (discharged) ions, excluding the adsorbed ones, are uniformly distributed throughout the dispersion medium, including the region between the plates. In the model of two infinite plates, however, when $\lambda \rightarrow 0$, the quantity $\lambda e_1\overline{\psi}(\lambda, x)/kT$ does not tend to zero as it should. It remains finite, reaching a value $8\pi e_1{}^2\nu_1/DkT\kappa_0{}^2R$ which depends on R. According to the Boltzmann distribution law, this means that the density of ions between the plates is different from that in the bulk of the solution and furthermore is a function of the separation.

One could calculate an osmotic contribution to the free energy associated with this hypothetical density of discharged ions between the two plates. This would have the form

$$\lim_{\lambda \rightarrow 1} ARkT \sum_{i=1}^{s} \overline{n}_i(\lambda, x) \ln \overline{n}_i(\lambda, x)$$

For small potentials, it is possible to show that this quantity yields the following contribution which depends on R,

$$\lim_{\lambda \rightarrow 0} \frac{D}{8\pi} AR\lambda^2\kappa_0{}^2\overline{\psi}^2(\lambda, x) = \frac{8\pi A e_1{}^2\nu_1{}^2}{D\kappa_0{}^2R} = \lim_{\lambda \rightarrow 0} 2A\int_0^{\sigma_0}\psi d\sigma.$$

---

[14] Palmer and Schmitt, *J. Cell. Comp. Physiol.*, 1941, 17, 385 (fig. 2).

The last expression is precisely the correction term introduced by Dr. Overbeek, the form given here is identical with that of the excess osmotic pressure term introduced by Derjaguin and is consistent with Langmuir's use of the osmotic forces.

Dr. Overbeek is quite correct in stating that such a term cannot be derived by the method described in my theory, which does not allow for such an effect. My system is defined by the three thermodynamic variables λ, ν_1 and R. In the discharged state, when only short-range forces are present, thermodynamic equilibrium of such a system implies a uniform distribution of ions. Essentially, the passage from this distribution to one having a density difference between the region of the plates and the bulk of the solution, involves a thermodynamically irreversible process and hence is inconsistent with the basis of my method.

It is possible to arrive at Dr. Overbeek's proposed correction in another way, which again presents difficulties. In my second paper (II) it is shown that if the potential function $\Phi_1(\lambda, R)$, which is defined as the potential of the mean force at the particle surface, becomes identical with the mean potential, and if the linear Debye-Hückel equation is used, then my electrical free energy function can be written as

$$2A\int_0^{\lambda}\frac{\overline{E}(\lambda,\ \nu_1,\ R)}{\lambda}d\lambda = A\lambda e_1\nu_1(\lambda_1 R)\overline{\psi}(\lambda,\ R/2)$$

where $\overline{\psi}(\lambda, R/2)$ is defined by equation (26) in my paper (I). It is assumed that the two plates are surrounded by an infinite volume of electrolyte so that the asterisk superscript is dropped. When $\lambda \rightarrow 0$ this energy becomes the limiting correction introduced by Dr. Overbeek. It follows from my equation (16) that the part of the mutual free energy depending on the separation is

$$-A\bar{\nu}_1(\lambda,\ R)[\lambda e_1\overline{\psi}(\lambda,\ R/2) + 2kT]$$

where we substitute $\nu_1 = \bar{\nu}_1(\lambda, R)$ into the formula (26) for $\overline{\psi}(\lambda, R/2)$. The function $\bar{\nu}_2(\lambda, R)$ is determined from the equilibrium equation (15) which now reads

$$\Delta\chi_1(\lambda) + \mathrm{B} - kT - kT\ln\bar{\nu}_1(\lambda,\ R) = \frac{4\pi\lambda e_1^2}{D\kappa_0}\bar{\nu}_1(\lambda,\ R)\coth\tfrac{1}{2}\lambda\kappa_0 R.$$

Noting that the formula (17*i*) now becomes

$$G(\lambda,\ R) = \lambda e_1\overline{\psi}(\lambda,\ R/2)/\bar{\nu}_1(\lambda,\ R)$$

it follows that the formula (13*i*) for the force is replaced by

$$-2A\bar{\nu}_1^2(\lambda,\ R)\,\frac{4\pi\lambda e_1^2}{D\kappa_0}\,\frac{\mathrm{d}}{\mathrm{d}R}(\coth\ \lambda\kappa_0 R)$$

which yields a repulsion. The objection to this analysis is that when $\lambda \rightarrow 0$ the surface density of ions $\bar{\nu}_1(\lambda, R)$ remains dependent on the separation R. However, the charging conditions at the start ($\lambda = 0$) must be independent of R.

A necessary condition for the validity of my method is that the product of λ and the potential in the double layers tends to zero as $\lambda \rightarrow 0$, assuming a fixed surface density of ions. This condition ensures that the mutual free energy be independent of the separation at the beginning of the charging process. However, it is not satisfied by the solution of the Poisson-Boltzmann equation for the potential in the region between two infinite plates. Consequently, we cannot treat as valid the particular application of my theory to two infinite parallel plates, given in my paper. At the same time, the proposed correction of Dr. Overbeek is also unsatisfactory.

In my second paper (II), I discuss the cases of two spherical and parallel cylindrical particles. It is shown that as the ions are discharged, at constant density of adsorbed ions, the potential in the double layers approaches zero and hence my theory can be applied. If the mean potential is identified with the potential of the mean force and the linear Debye-Hückel equation is used, then we obtain an expression for the free energy which has the same form as that given above. The quantity $A(\bar{\nu}_1\lambda, R)$ is replaced by the total number of absorbed ions on each particle under equilibrium conditions, and in place of $\bar{\psi}(\lambda, R/2)$, we introduce the average potential over each particle surface. Under these assumptions, a repulsion is again obtained but the formula for the free energy is different from that proposed by Drs. Verwey and Overbeek.

If the edge effects in the case of two parallel plates of finite area are introduced, then we can expect the quantity $\lambda\bar{\psi}(\lambda, x)$ to vanish when $\lambda \rightarrow 0$. However, the analysis in my second paper suggests that on the basis of the approximate Debye-Huckel equation for the potential, a repulsion will be obtained. Such a result is at first sight rather unexpected when the dimensions of the plates are large compared with their separation, for then the correction for the edge effects should be small although, sufficient to prevent the potential from rising to infinity as $\lambda \rightarrow 0$. It should be remarked that an attraction between two such plates implies that the mean potential is different from the potential of the mean force.

Drs. Verwey and Overbeek avoid the difficulty of the infinite potential by assuming that the surface potential is kept constant. However, it is not all clear that this assumption is consistent with the necessary condition that at each stage of the charging process, we have true equilibrium. Furthermore, my method shows that the surface potential is a function of the particle separation.

The problem of ensuring equilibrium conditions during the charging process can be clarified by some remarks on the work of Derjaguin. This author separates the free energy into a potential energy term and an osmotic term. The former is the work done in charging the ions and colloidal particles assuming that the ions are fixed in their final average positions. The latter represents the work done against the thermal motion resulting from the re-distribution of the ions during the charging process. His approach is equivalent to assuming that the discharged ions first re-distribute themselves so as to be in their final average positions and then they are charged. The objection to this method is that strictly it is not an isothermal reversible process. Nevertheless, it can be shown that Derjaguin's formula for the free energy is identical with my expression for the electrical free energy provided (i) the mean potential is identified with the potential of the mean force, (ii) the potential energy is a quadratic function of the density of adsorbed ions, which is the case with the approximate Debye-Hückel equation, (iii) the surface density of ions is independent of particle separation, so that the " chemical " free energy is absent.

I agree with Dr. Overbeek that the Poisson-Boltzmann equation is probably a good approximation and that the fact that $\bar{E}_0(\lambda, R) - \bar{E}_0(\lambda, \infty)$ is negative does not imply an attraction between the particles. However, it has not been definitely proved that a repulsion always exists between the double layers of colloidal particles. Firstly, the case of large surface potentials remains to be examined by my method. Secondly, the forces responsible for tactoid formation are quite small and it is not certain that the Poisson-Boltzmann equation always gives the correct result.

Prof. J. D. Bernal (*London*) said : The controversy between different formulations of long range forces indicates that we are still some way from a fundamental understanding of the nature of these forces. All theories propounded so far only claim to give a general account of the equilibria which may exist between particles in a solution, and the relation of these equilibrium distances to conditions in the medium. They cannot at this stage be effectively tested by numerical comparison with observed results. This is very largely because the experimental evidence itself is patchy and imperfect. None of the systems studied can be adequately characterized either as to the size of the particle or its surface charge. Most of them, such as vanadium pentoxide and hydroxide or clay suspensions, are poly-disperse, and on tobacco mosaic virus, which is the most uniform, the physical measurements were only of a preliminary kind. It is clearly essential that we get more experimental evidence in order to arrive at a reasonable model on which to base calculations.

What strikes me in the present stage of investigation, however, is the extremely general occurrence of long distance equilibria, and this makes me speculate that we may have here some phenomena which do not depend on particular features of surface charge and potential, or on the precise distribution of ions in the solution. Is there not a possibility that in any heterogenous mixture, for instance of two substances of different dielectric constants, the free energy is a minimum if the substance is divided into layers of equal thickness for each phase ? Is there not a possibility of a more general London type of mutual interaction of polarisable phases which would give such a result ?

A METHOD FOR MEASURING THE TRANSVERSE SWELLING OF FIBRES.

By J. M. Preston.

Received 19th August, 1946.

Textile fibres swell when placed in water. The way they swell depends on their chemical constitutions and crystallinity, and affects their behaviour during dyeing and finishing. For these reasons a knowledge of their swelling behaviour would be valuable but has proved hard to ascertain because of their small dimensions. Few textile fibres are shorter than about an inch in length whilst many are much longer. Few fibres are more than one thousandth of an inch in breadth and many are smaller. These dimensions make it comparatively easy to measure lengthways swelling and difficult to measure transverse swelling. Furthermore fibres are anisotropic therefore the transverse swelling cannot be directly inferred from the lengthways value. If the swelling values could be determined in the two directions their ratio would be a valuable index of their anisotropy.[1]

A number of investigators have measured the transverse swelling of fibres. Mostly the microscope has been used [2, 3, 4, 5, 6, 7, 8], but only under the very best conditions do these measurements show reasonable accuracy. Some of the purely microscopical difficulties have been recorded [9,10,11], and there is no need to discuss them here. There is one difficulty peculiar to transverse swelling measurements under the microscope which is worth mentioning. This is caused by the irregular outlines of the transverse sections. On swelling the fibres tend to revert to the original form in which the sections were more nearly circular. To do this the greater diameters swell less than the smaller ones. Breadth measurements are made automatically of the outlines of the fibres. These increase less on swelling than the bases of the re-entrant angles and other less easily observed smaller diameters. Therefore breadthways swelling measurements made with a microscope tend to be minimum values. Better values for the mean transverse swelling can be obtained from measurements of cross-sections, but the method is very laborious.

---

[1] Hermans and de Leeuw, *Kolloid Z.*, 1937, **86**, 300.
[2] Lawrie, *J. Soc. Dyers Col.*, 1928, **44**, 73.
[3] Clayton and Peirce, *J. Text. Inst.*, 1929, **20**, T315.
[4] Speakman, *Trans. Faraday Soc.*, 1929, **25**, 95.
[5] Collins, *J. Text. Inst.*, 1930, **21**, T311.
[6] Denham and Dickinson, *Trans. Faraday Soc.*, 1933, **29**, 300.
[7] Steinberger, *Text. Research*, 1933/4, **4**, 33.
[8] Macmillan, *J. Text. Inst.*, 1939, **30**, P305, T73.
[9] Quastler, *Z. Wiss. Mik.*, 1932, **49**, 195.
[10] Bishop, *Bur. Stand. J. Res.*, 1934, **12**, 173.
[11] Cassie, *J. Text. Inst.*, 1942, **33**, T195.

Other methods have been tried for measuring swelling. In one a volumenometer has been used.[12] This suffers from the difficulty of deciding when to stop the centrifuging which precedes the volume measurement. There is no-end point to this since it is found that the amount of moisture retained in fibres is a continuous function of the speed and duration of centrifuging.[13, 14] The suggestion was also made that air permeability measurements might be used,[11] but the method proved unsuitable.[15]

Some years ago Viviani [16] described a method for recording variations in yarn thickness from place to place in a yarn. The method depended on the displacement of mercury from a capillary tube by the yarn. The electrical resistance of the tube was an index of the amount of mercury displaced. By a purely empirical calibration this resistance could be correlated with the denier of any particular yarn. The method is very useful for tracing yarn variations but does not give a measurement of the cross-sectional area of the fibres because of the high surface tension of the mercury. The principle of this method could be applied to swelling measurements if a suitable working fluid could be substituted for mercury.

The characteristics required in the working fluid for swelling measurements are that:

(i) it should leave no empty spaces around or between the fibres;
(ii) it should have a high conductivity compared to that of the swollen fibres;
(iii) it should swell fibres in a similar way to water.

These requirements would be met by an electrolyte solution in which the fibres swelled as in water and which did not confer appreciable conductivity on the fibres.

The conduction of electricity through moist fibres is considered to be electrolytic and to take place through the pores[17, 18]. The pores in wet fibres have been estimated to be of the order of 20-40 A. in diameter. [19, 20, 21]. In considering the electrical resistance of moist fibres O'Sullivan [22] has suggested that the proximity of the walls of the pores increases the viscous hindrance to the movement of the hydrated ions. This effect is like the sieve effect which has been observed with membranes.[23] It is similar to the wall effect present in falling sphere viscometers. Francis has given an expression, $(1 - R/r)^{2 \cdot 25}$, for the retarding effect of the walls in the macroscopic case.[24] In this expression R/r is the ratio of the diameter of the ball to that of the tube. If these considerations apply at the dimensions of the pores then the electrolytic conduction through a fibre should diminish rapidly as the ion size approaches that of the pore. Ions

[12] Eckert and Hoinkis, *Papierfrabicant*, 1937, **35**, 441.
[13] Coward and Spencer, *J. Text. Inst.*, 1923, **14**, T28.
[14] Bancroft and Calkin, *Text. Research.*, 1933/4, **4**, 371.
[15] Cassie (private communication).
[16] Viviani, *Kunstseide*, 1933, **15**, 30.
[17] Murphy and Walker, *J. Physic. Chem.*, 1928, **32**, 1761.
[18] Marsh and Kathleen Earp, *Trans. Faraday Soc.*, 1933, **29**, 173.
[19] Morton, *ibid.*, 1935, **31**, 262.
[20] Lenher and Smith, *Ind. Eng. Chem.*, 1935, **27**, 20.
[21] Meitner, *J. Soc. Dyers Col.*, 1945, **61**, 33.
[22] O'Sullivan, *Trans. Faraday Soc.*, 1933, **32**, 192.
[23] Meyer and Sievers, *Helv. Chim. Acta.*, 1936, **19**, 649, 665.
[24] Francis, *Physics*, 1933, **4**, 403.

of the largest effective size should be the best for measuring the swollen dimensions of fibres.

A number of estimates have been published of the hydration of ions. In a recent paper[25] Darmois has given the ionic radii. The following values are selected from his tables: K 1·19 A., Li 2·31 A., Be 5·52 A., SO_4 2·24 A. If cations and anions have to pass each other in a pore then the sum of their radii must be less than that of a pore for conduction to take place. The sum of the radii of the beryllium and sulphate ions approaches that estimated for the pore size in fibres. On this hypothesis fibres soaked in beryllium sulphate solution should conduct less than when lithium sulphate is used and the latter should conduct less than when potassium sulphate is the electrolyte.

The ideal electrolyte should confer no appreciable conductivity on the fibres. In this case the conductivity in the capillary tube will be directly proportional to the area not occupied by fibres. If there is appreciable conduction through the fibres this simple relation will not hold. In the general case let the conductivity of the cell containing the capillary filled with electrolyte alone be C_0 and when filled with fibres and electrolyte be C_f. Let the relative conductivity of the swollen fibres be k times that of an equal cross-section of the solution and let α be the fraction of the capillary filled by the fibres. Then

$$C_f = \{ (1 - \alpha) + \alpha k \} . C_0.$$

Therefore

$$\alpha = \frac{1 - C_f/C_0.}{1 - k}$$

The values of k for different electrolytes can be found by comparing the *cross-sectional areas of fibres measured by the conductivity method* and some independent method. The smaller is k the more accurate will be the conductivity method. The only independent method is the microscopical one. If very large fibres are used this can have an accuracy of measurement of about 1 in 200. In addition to being large the fibres should be cylindrical to avoid complications in measurements and calculations.

The third characteristic required of the conducting fluid is that the fibres shall be swollen in it to the same extent as in water. This may not be completely fulfilled by the electrolytes discussed, but the defect may be small. There is an indication that this is probable in the observations that the weights of salt solutions taken up by Cellophane are in proportion to their densities.[26]

In order to determine the transverse swelling of fibres it is necessary to know their transverse dimensions dry as well as wet. The dry value is derived simply from the specific volume which is easily determined with anhydrous fibres by pyknometric or flotation methods.[27] Data for all the common fibres are given in the literature.

Apparatus.—The conductivity cell used for measuring the cross-sectional areas of the swollen fibres is shown in Fig. 1. It is very simple and consists of two similar J-shaped resistance glass tubes about 20 cm. high and 2 cm. diameter. These are joined together by rubber stoppers and a piece of capillary tube about 7·5 cm. long and 0·180 cm. internal diameter. A pair of similar cells is used immersed in a thermostat controlled to ± 0·1° C. One of the cells acts as a control and is filled with the electrolyte alone. In the other cell a known weight of dry fibres of known length is threaded through the capillary. The resistance of the capillary is so large compared to the other resistances in the cell that the others can be neglected.

The conductivity of the cell could be measured by the ordinary Wheatstone bridge circuit, but it was thought that a direct reading circuit would have advantages. The conductivity was read directly with the use of stepped-down and stabilised 230 volts, 50 ~ A.C. mains current. The stabiliser and the rest of the circuit are shown in Fig. 2. In this T_1 is a 230 v. : 75 v. transformer and T_2 is a 230 v. : 36, 30, 24, 18 and 12 v. tapped transformer. The appropriate tapping is selected depending on the conductivity of the cell. R and R are 960 Ω Constantan wire wound resistances. L and L are 40 w. 230 v. gas-filled coil lamps The wire wound resistances are immersed in an oil bath to maintain an approximately constant temperature. The lamps should be kept free from vibrations. These disturb the filaments and cause fluctuations in the voltage. M is a rectifying microammeter and C is the conductivity cell.

---

[25] Darmois, *J. Chim. Physique*, 1946, **43**, 1.

[26] Usher and Wahbi, *J. Soc. Dyers Col.*, 1942, **58**, 281.

[27] Hermans, *A Contribution to the Physics of Cellulose Fibres* (Elsevier, 1946), p. 207.

The stabilising circuit was tested with a wire wound resistance substituted for the cell. The resistance was adjusted to give nearly full scale deflection of the meter and then the input voltage was changed with a Variac It was found that a variation of the input of either + or − 30 v. about a mean value of 210 v. produced a decrease of meter reading of only ½ % in each case. In another test the circuit was run continuously for six hours without any change of reading being noticeable after the first five minutes during which the circuit warmed up. The linearity of the meter readings was tested with wire wound resistances substituted for the cell. The readings showed only very slight deviations from a straight line when plotted against the known conductivities. To get the greatest accuracy a calibration curve was constructed and the conductivities were read from this.

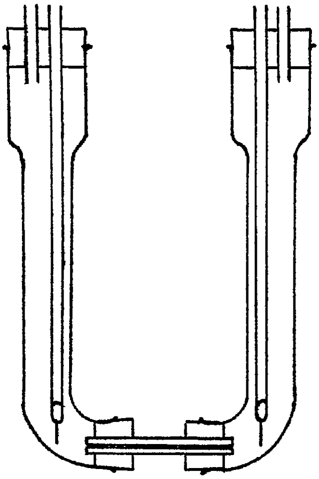


FIG. 1.

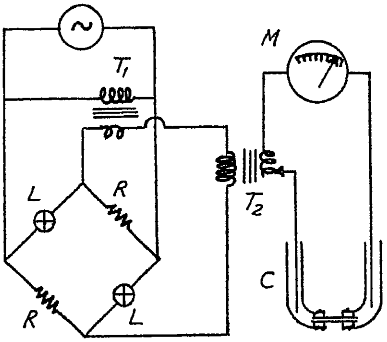


FIG. 2.

Method.—The dimensions of the swollen fibres are obtained by difference between the readings of the conductivities of the cell with and without fibres in it. The greater the proportion of the capillary filled with fibres the more accurately can this be measured, but there is an upper limit. This is set by the need not to impose any restraint on the swelling of the fibres. By trials with cuprammonium monofils which had swollen diameters ¼ that of the capillary, it was found that not more than eight monofils could be packed in without their pressing on each other in the swollen condition. This filled only ⅓ of the cross-sectional area. Geometrically it should have been possible to have fitted ten monofils in if these had been perfectly packed. When ten were used a low swelling value was obtained and the swollen monofils were difficult to withdraw from the capillary. If the fibres are very fine, cylindrical and packed hexagonally it should be possible to fill about 90 % of the capillary. In practice it is safest not to attempt to fill more than half the capillary with fibres.

The bundle of fibres is drawn into the capillary dry. The capillary is next connected to the J-shaped tubes. Electrolyte solution is poured into one of the tubes and the solution is allowed to irrigate the fibres thoroughly as it flows through the capillary. Both sides of the cell are connected to a water pump for a short time to remove air bubbles and placed in the thermostat. After about half an hour the conductivity is read on the meter. Later the meter is read again after tilting the cell to irrigate the fibres to check that equilibrium has been obtained. There is no change of reading after 15 min. with viscose

---

[18] Glynne, *J. Inst. Elec. Eng.*, 1943, 90, Pt. II, 101.

rayon fibres, but because of their great thickness 24 hours were allowed before taking readings in the case of the monofils

Results.—To obtain a measure of the accuracy of the conductivity method comparison was made with the microscopical method. To obtain reasonable accuracy with the latter, large cuprammonium and Nylon monofils were used as test fibres. The cuprammonium and the larger of the Nylon monofils were big enough to reduce the refraction and diffraction effects to negligible values. The errors from these causes with the finer Nylon monofils probably did not exceed 1 %.

Comparisons were also made of the diameters of the monofils in water and various salt solutions. The difference was less than ½ % for the cuprammonium monofils and less than 1 % for the smaller Nylon monofils. The differences were within the experimental errors.

The diameters of the monofils were calculated from the conductivity measurements in various salt solutions. It was assumed that there was no conduction through the fibres in making the calculations (i.e. $k = 0$). The results are set out in Table I.

TABLE I

Fibres.	Electrolyte.		*p*H.	Fibre Diameters.		Fibre Conductivity. k
	Concn.	Salt.		Microscope.	Conductivity.	
				mm.	mm.	
Cuprammonium	M/2	K_2SO_4	—	0·307	0·276	0·19
	,,	Li_2SO_4	—	0·307	0·290	0·11
	M	$LiOOC.CH_3$	—	0·307	0·288	0·12
	M/2	$BeSO_4$	2·2	0·307	0·302	0·03
Nylon . .	,,	Li_2SO_4	—	0·059	0·058	0·03

The measurements were made at 25° c. with the conductivity apparatus and at 20° c. with the microscope.

From a comparison of the fibre diameters it is obvious that there was some conduction through the fibres in all cases. In no case was $k = 0$. The calculated values of k are set out in the last column. The three cations show a definite trend with the cuprammonium monofils; k tends towards zero as the effective size of the ions increase. The difference of the diameters in the case of the Nylon monofils is only slightly larger than the probable error of microscopic measurement.

Beryllium sulphate gave the best results but had two defects. One defect was a drift of the conductivity with time. This was due to complex formation. It could be speeded up and brought to finality by boiling for a few minutes. The other defect was the comparatively high acidity of the solution. This had a *p*H of 2·2 when freshly prepared. At this *p*H the hydrogen ion concentration is not negligible and must contribute an appreciable fraction of the total conductivity of the solution. Hydrogen ions are known not to be retarded in the pores of membranes as are larger ions.[23] If the pores in the fibres permitted the hydrogen ions to show their normal high conductivity it would account for the value of k found even if the beryllium and sulphate ions had no conductivity through the pores. If this explanation is correct the conductivity through the cuprammonium fibres should be appreciably reduced by decreasing the hydrogen ion concentration in the beryllium sulphate solution. To do this the complex basic beryllium sulphate was prepared by boiling a M/2 solution of the sulphate and adding beryllium carbonate at the rate of 33 g. for each litre of solution. This brought the *p*H to 4·4. At this value the hydrogen ion concentration should contribute a negligible fraction of the conductivity of the M/2 sulphate solution.

A comparison was made of the diameters of the cuprammonium and Nylon monofils in the basic beryllium sulphate solution measured with a microscope and the conductivity apparatus. The results are set out below in Table II. The two sets of measurements agree within the accuracies of measurement which were about 1 in 200 for diameter measurements and 1 in 100 for area measurements. This shows the conductivities though the pores of the fibres were negligible.

TABLE II

Fibres.	Electrolyte.		pH	Fibre Diameters.		Fibre Conductivity. k
	Concn.	Salt.		Microscope.	Conductivity.	
				mm.	mm.	
Cuprammonium	M/2	$BeSO_4$	4·4	0·274$_4$	0·275	0·00
Nylon . .	,,	,,	4·4	0·216 0·208	0·214 0·209	0·01

A further check on the accuracy of the conductivity method was made with two viscose rayons. These had previously been measured very carefully on photomicrographs of their cross-sections. The identical sections were measured wet and dry. The ratios of these areas gave the transverse area swelling. The same ratios were determined by the conductivity method using basic beryllium sulphate solution. The results are given in Table III. The swelling ratios were calculated on the completely dry condition. For this purpose a correction was applied to the air-dry microscopical measurements. A correction was also applied to allow for the axial swelling in calculating the corresponding cross-sections in the conductivity method. These corrections were derived from data given by Hermans [2] and from the experimentally determined axial swelling values. The lower swelling ratios found by the microscopic method are possibly due to the systematic displacement of microscopical measurements reported by Bishop [10] and Herzog.[29]

TABLE III

Fibre.	Transverse Area Swelling.	
	Microscope.	Conductivity.
Viscose stretched 0 % .	1·99	2·14
Viscose stretched 30 % .	2·07	2·05

With large fibres there is probably nothing to choose between the microscopical and conductivity methods for accuracy, but the latter is much less laborious. For fine fibres the conductivity method is both quicker and more accurate.

The investigation described in this paper was done at the request of Messrs Courtaulds Ltd. who have kindly given permission for its publication, and to whom the author wishes to acknowledge his thanks. He also wishes to thank Dr. R. V. Bhat and Mr. P. C. Tsien who carried out measurements included in this paper.

Dept. of Textile Chemistry,
College of Technology,
Manchester 1.

[29] Herzog, *Die mikroskopische Untersuchung der Seide und der Kunstseide* (Springer, 1924), p. 32.

SORPTION, SWELLING AND ELASTIC CONSTANTS OF THE CELL WALL MATERIAL IN WOOD.

By Wilfred W. Barkas.

Received 18th July, 1946.

This paper shows how the sorption and swelling of hygroscopic materials may be used to give quantitative information about their elastic properties. The fundamental theory will be equally applicable to materials which absorb vapours other than water, but the applications will be confined to the sorption of water by wood, for which purpose the subject was originally developed.

In Part I is given the mathematical theory of swelling stresses and the calculation of elastic constants.* The equations derived in this Section are in differential form and are exact. Mathematical integration is not attempted, the integral values being obtained by the summation of small steps, each of which is based on empirical data. In Part II the use of the theory is illustrated by the calculation of elastic constants of the natural wood and of the wood substance composing the cell walls. Finally, Part III contains a discussion of certain consequences of the theory of swelling stresses such as sorption hysteresis, the instability of capillary spaces on drying, and the limiting value of the adsorption moisture content at the saturation vapour pressure.

Part I. Theory.

Empirical Data.

There are several generally accepted observations on the swelling and sorption of gels. These include : (1) gels sorb moisture and hold it at a reduced vapour pressure ; (2) on sorbing moisture, they change in dimensions ; (3) they possess rigidity ; (4) they exert directional stresses on an elastic enclosure ; (5) they show limited swelling ; and (6) they show hysteresis in the sorption isothermal. The hysteresis loop does not necessarily close at a saturated vapour pressure [2] (Fig. 1). The exact theory of swelling stresses will be based on observations (1) to (3) and will require two additional assumptions : (1) that the only volatile constituent is water, and (2) that the gel is perfectly elastic.

Differential Swelling.—If a block of gel measures x cm. in thickness and $y \times z$ in cross-section, and if it sorbs ∂m g. of moisture, the differential swellings are

$$\left.\begin{aligned} s_x &= \frac{V}{x}\left(\frac{\partial x}{\partial m}\right)_{XYZ} \\ s_y &= \frac{V}{y}\left(\frac{\partial y}{\partial m}\right)_{XYZ} \\ &= \frac{V}{z}\left(\frac{\partial z}{\partial m}\right)_{XYZ} \end{aligned}\right\} \qquad (1)$$

and

$$s_{XYZ} = s_x + s_y + s_z, \quad . \quad . \quad . \quad . \quad (2)$$

---

* This is largely summarised from the author's report on "Swelling Stresses in Gels" [1] referred to in what follows as "the Report."

[1] Barkas, *Swelling Stresses in Gels*, Forest Products Research Special Report No. 6, H.M.S.O. (1945.)

[2] Barkas, *Proc. Physic. Soc.*, 1936, 48, 576.

where V is the volume of gel at moisture content m which contains 1 g. of dry material and where in all cases, X, Y, and Z must be kept constant during the measurement of swelling.

The following equations (3) to (9) make use of purely empirical measurements and are independent of any theory of the internal structure or hygroscopic behaviour of the gel. The derived stresses and moduli apply to any conditions identical with those under which the s terms are measured.

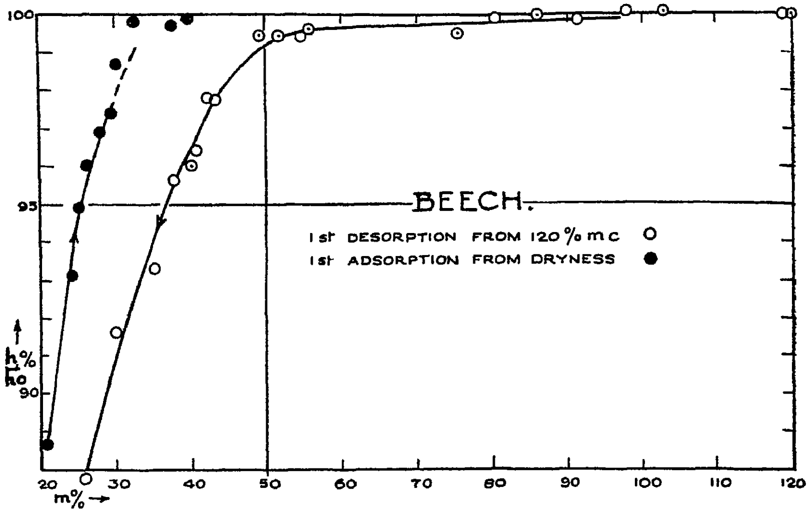


FIG. 1.

Stress.—If the stress-free differential sorption $(\partial h/\partial m)_{XYZ}$ be written D, and that under a restraint, D', then for a gel adsorping dm' g. of moisture under perpendicular stresses, X, Y and Z we have

$$s_x\frac{dX}{dm'} + s_y\frac{dY}{dm'} + s_z\frac{dZ}{dm'} = v(D' - D), \quad . \quad . \quad . \quad (3)$$

where v is the specific volume of water vapour at pressure h.

Other geometrical data are required to enable one to calculate X, Y and Z separately from (3) though it may be used to calculate dm' from a given set of stresses. For a finite rise in vapour pressure and a finite sorption, the s terms must be measured over a series of small steps since s and D are seldom known functions of h.

Strains.—For a gel sorbing dm' g. of moisture, the elastic strains parallel to the x, y and z axes are

$$\frac{(s_x' - s_x)}{V}dm'; \quad \frac{(s_y' - s_y)}{V}dm' \quad \text{and} \quad \frac{(s_z' - s_z)}{V}dm'.$$

These may be used in the usual equations for bodies having rhombic symmetry.

Bulk modulus, if $X = Y = Z = p$,

$$k = \frac{Vv}{s_p}\left(\frac{D' - D}{s_p - s'}\right). \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Young's modulus, if $Y = Z = 0$,

$$E_x = \frac{Vv}{s_x}\left(\frac{D' - D}{s_x - s'_x}\right)_{Y=Z=0} \quad . \quad . \quad . \quad . \quad (5)$$

and similarly for E_y and E_z.

Poisson's ratio, if $Y = Z = 0$.

$$\sigma_{xy} = -\left(\frac{s'_y - s_y}{s'_x - s_x}\right)_{Y=Z=0} \quad . \quad . \quad . \quad (6)$$

and similarly for the other five ratios.

If we confine ourselves to stresses acting parallel to the principal axes, the rigidity n is calculated from (5) and (6) in the usual way. If necessary, the other strains needed to complete the analysis of the rhombic system can be measured, in which case three more *independent* rigidity moduli appear.

If a second system of stress is applied, denoted by a double dash, we have also as an alternative to (4),

$$k = \frac{Vv}{s_p}\left(\frac{D'' - D'}{s' - s''}\right) \quad . \quad . \quad . \quad . \quad (7)$$

and similarly for E_x, etc.

No stresses need be measured directly since these are obtained from the reduction in sorption by equations similar to (3).

Constant Vapour Pressure.—The differential swelling under stress at constant vapour pressure is $(\partial V/\partial m)_h = s_h$ and D' is zero, so equations (4) or (7) become

$$k = \frac{VvD}{s_p(s_h - s_p)} \text{ or } \frac{VvD''}{s_p(s_h - s'')} \quad . \quad . \quad . \quad . \quad (8)$$

and s_h may be calculated from (4) and (8), or (7) and (8)

$$s_h = \frac{s_pD' - s'D}{D' - D} \text{ or } = \frac{s'D'' - s''D'}{D'' - D'} \quad . \quad . \quad (9)$$

From calculations made on solutions of known bulk modulus, s_h is found to be nearly constant for all solutions and it will be approximately applicable to non-porous, but not to porous, gels. Values of s_h and other data on the non-porous cell-wall material of spruce wood are given in Table I.

TABLE I.—SPRUCE WOOD SUBSTANCE ADSORPTION DATA AT 25° C.

R.H. %	100	95	90	80	60	40	20	10
Moisture content : (m') – –	0·30	0·254	0·209	0·163	0·114	0·084	0·054	0·035
Volume per g. dry weight : (V_3) –	—	0·906	0·861	0·817	0·770	0·774	0·719	0·705
Differential swelling :—								
at constant stress : (s'_3) –	—	0·991	0·980	0·962	0·917	0·872	0·788	0·673
at constant vapour pressure : $(s_3)_h$ – – – – –	1·003	1·005	1·009	1·016	1·036	1·065	1·116	1·166

Part II. Applications to Wood.

Cell Wall Material as a Gel Structure.

Natural wood is a porous material composed of hollow thick-walled cells whose length is considerable in comparison to their diameter and which are easily visible under the microscope. The cells fit together externally so that the only voids in the structure are those of the cell lumen. The swelling of the cell occurs principally in the transverse plane and the longitudinal swelling may, for present purposes, be considered negligible. Though, for anatomical reasons, the swellings of the gross wood in the radial and tangential directions perpendicular to the grain are unequal,[3] there is no reason to doubt that the transverse swelling of the gel composing wood substance is isotropic. Thus, for purposes of calculations I propose to introduce what is perhaps an over-simplified model of the unit cell, regarding it as a long hollow tube of hygroscopic material with zero longitudinal swelling and of isotropic transverse properties, thus reducing the calculation to a problem of two-dimensional elasticity.

---

[3] Barkas, *Trans. Faraday Soc.*, 1941, **37**, 535.

Assuming that the wood is composed of uniform cells, possessing the average properties of real cells, it will be found that the absolute dimensions of the unit cell do not appear in the equations, so that we may make our unit cell of external radius r_1 and internal radius r_2 such that 1 cm. length contains a volume V of gel having 1 g. dry weight and that, therefore, the weight of water in it is numerically equal to the moisture content.

In this way let V_1 be the external volume, V_2 that of the hollow lumen and $V_3 = V_1 - V_2$, that of the wood substance. In a sample of wood, V_1, V_2 and V_3 will refer respectively to the external volume, the void volume, and the gel volume, per g. dry weight. Of these, V_1, which varies considerably in different samples, may easily be measured directly at all moisture contents, and V_3, which appears to be reasonably constant for all samples, has been determined.[4] Thus the variation in density between different samples of the same species is mainly a question of variations in the proportion of void space, known as the cell space ratio, A.

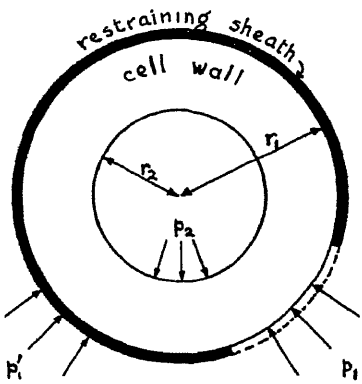


FIG. 2.

The corresponding s terms are the differentials of these volumes with respect to moisture content. When these differential swellings are compared, it is a simple matter to show that there must be some restraint imposed on the free-swelling of the wood substance. Anticipating this proof, I shall, therefore, adopt the previous convention of using dashes when referring to swelling under stress. Thus the natural differential swelling of the gross wood, though apparently free from stress from the point of view of externals, is really occurring under restraint and will therefore be denoted by $s_1' = dV_1/dm'$ and that of the wood substance by $s'_3 = dV_3/dm'$. The elastic constants will likewise be denoted by k', E', etc.

In the absence of restraint, s_1 would be $s_3 V_1/V_3$ and the volume increase of the hollow centre s_2 would be $s_3 V_2/V_3$, just as in the well-known method of calculating the thermal expansion of a hollow glass vessel from that of the glass itself. In the case of wood, however, it is found that s_1' is only about one-third of this calculated value, and that s_2' is so small that the cross-section of the lumen can be taken as being approximately constant at all moisture contents. Evidently, then, some restraint is operative and we are justified in using dashes to designate the natural hygroscopic properties of wood.

We may now enquire into the source of this natural restraint on the swelling of the wood substance. When wood sections observed under a microscope are treated with dilute alkali which causes excessive swelling, it is found that initially the greatest swelling occurs *into* the lumen, which may completely disappear, and that only later does the outer wall of the cell disrupt, whereupon the wood substance disperses into a much larger volume. From this we may conclude that the principal restraint to swelling lies in the outer periphery of the cell wall and that, if there is an internal restraint next the lumen, it is much less effective than the other.

---

[4] Stamm and Seborg, *J. Physic. Chem.*, 1934, **39**, 133.

Thus it is now permissible to elaborate the original model of the unit cell to include a non-swelling but elastic outer sheath surrounding the gel material of the cell wall. This sheath expands under internal pressure but does not swell hygroscopically and may be taken to have negligible thickness compared to that of the cell wall itself, so we may continue to speak of the gel material as the " wood substance " without introducing serious error. There is experimental evidence for the anatomical existence of this restraining outer wall, and it is believed to consist of a flat spiral winding of cellulose fibrils.

Since the outer sheath cannot be removed experimentally we must find some way of circumventing the effects of the natural restraint by theoretical means, if we are to be able to calculate the elastic properties of the stress-free wood substance. Obeying the above convention, plain symbols s_3, k_3, etc. will be used to designate the hygroscopic and elastic properties of this unstressed material.*

Similarly, p_1 will be taken as the radial pressure exerted externally on the hollow tube of wood substance and p_1' as the radial pressure exerted on the restraining sheath by any forces acting externally on the gross wood. Thus in natural swelling p_1' will be zero (or, more strictly, equal to the vapour pressure, which is usually negligible) and p_1 will be the pressure exerted on the wood substance by the sheath itself (Fig. 2).

If under natural swelling, dm' g. of water are adsorbed, the compressive strains imposed on the hollow cell are :

$$\left.\begin{array}{ll} \text{for the cell} & \left\{\dfrac{s_1'}{V_1} - \dfrac{s_3}{V_3}\right\}\mathrm{d}m' \\[2ex] \text{and for the wood substance} & \left\{\dfrac{s'_3}{V_3} - \dfrac{s_3}{V_3}\right\}\mathrm{d}m'. \end{array}\right\} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

From the two-dimensional elasticity of a hollow tube it may be shown [1] that

$$\left.\begin{array}{ll} \text{for the cell} & \left\{\dfrac{s_1'}{V_1} - \dfrac{s_3}{V_3}\right\}\mathrm{d}m' = -\dfrac{\mathrm{d}p_1}{1 - A}\left\{\dfrac{1}{k_3} + \dfrac{A}{n_3}\right\} \\[2ex] \begin{array}{l}\text{and for the wood}\\ \text{substance}\end{array} & \left\{\dfrac{s_3'}{V_3} - \dfrac{s_3}{V_3}\right\}\mathrm{d}m' = -\dfrac{\mathrm{d}p_1}{1 - A}\,\dfrac{1}{k_3}, \end{array}\right\} \quad . \quad . \quad (11)$$

where A is the cell space ratio, i.e.

$$A = \frac{V_2}{V_1} = \frac{V_1 - V_3}{V_1}. \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Equations (11) give the ratio of k_3/n_3

$$\frac{k_3}{n_3} = \frac{s'_3 - (1 - A)s_1'}{A(s_3 - s'_3)} \quad . \quad . \quad . \quad . \quad . \quad (13a)$$

for natural swelling or, if external stresses are applied at constant humidity,

$$\frac{k_3}{n_3} = \frac{(s_3)_h - (1 - A)(s_1)_h}{A\{s_3 - (s_3)_h\}}. \quad . \quad . \quad . \quad . \quad (13b)$$

By analogy with (8), if external stresses are applied at constant humidity,

$$k_3 = \frac{V_3 v D'}{s_3\{(s_3)_h - s'_3\}}. \quad . \quad . \quad . \quad . \quad . \quad (14)$$

---

* It is, however, open to question as to whether the gel is really " stress-free " even when the restraint of the outer sheath of the cell has been removed, since it seems probable that it is the restraining action of occasional lateral linkages between molecules which prevents cellulose in water from dispersing into solution (see p. 149).

In order to calculate k_3, values of $(s_3)_h$ and s'_3 are taken from Table I, while s_3, the differential swelling of the wood substance when free of restraint is given by (13*a*) and (13*b*), i.e.

$$s_3 = \frac{V_3\{(s_1)_h s'_3 - (s_3)_h s_1'\}}{V_3\{(s_1)_h - s_1'\} - V_1\{(s_3)_h - s'_3\}}. \quad . \quad . \quad (15)$$

Equations (13), (14) and (15) are sufficient to solve for k_3 and n_3 and, since the wood substance is assumed to be two-dimensionally isotropic,

$$E_3 = \frac{4kn}{k + n}, \quad . \quad . \quad . \quad . \quad . \quad (16)$$

$$\sigma_3 = \frac{k - n}{k + n}. \quad . \quad . \quad . \quad . \quad . \quad (17)$$

This completes the equations for the elastic constants of the wood substance.

Experimental.—To test some of the above relationships, experiments on spruce wood were designed in which 2·5 cm. cubes were cut with the faces parallel to the principal planes and, after vacuum drying over P_2O_5, were conditioned to various humidities while the swelling was constrained by suitably designed clamps. The conditioning of such samples, even in vacuum, is a slow process and, owing to various accidents and delays, a complete set of reliable data is not yet available. The results obtained so far, however, are of sufficient interest to be reported here.

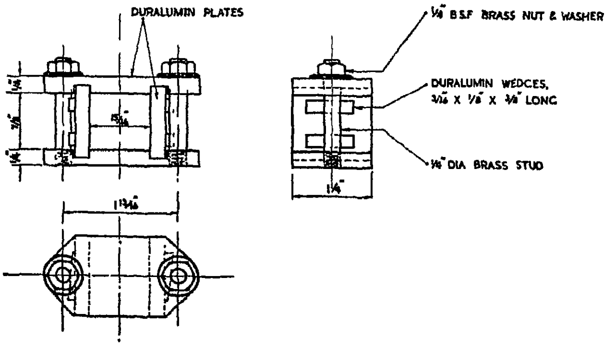

DURALUMIN PLATES
1/4" BSF BRASS NUT & WASHER
DURALUMIN WEDGES, 3/16" X 1/8" X 3/4" LONG
1/4" DIA BRASS STUD
3/16"
7/8"
1/4"
15/16"
1 1/4"
1 13/16"

FIG. 3.

Owing to the small longitudinal swelling of wood, little change in sorption results from clamps acting on the end grain faces, and there is also an advantage so far as conditioning is concerned, in leaving these free, since it is through the end grain that the most rapid penetration of moisture occurs. Light duralumin clamps were therefore designed as in Fig. 3, having brass bolts and nuts. Two side plates could be pressed up against the samples by wedges and, once fixed, they exerted a pressure during any subsequent swelling. Two matched controls were used with each clamped sample to give the stress-free behaviour of similar material.

The conditioning of the samples was done in a vacuum desiccator connected with a vessel containing an evacuated saturated solution of a suitable salt, and with a mercury manometer on which the vapour pressure could be measured directly. At each stage the samples were conditioned to constant weight in a thermostat at 23° C., ± 0·01°, and the three dimensions measured on a dial gauge. After being vacuum dried over P_2O_5 to obtain the dry weights, four samples were put through the following stages:

(A) Samples conditioned stress free in the lower humidity (32·8 °/₀ R.H.).

(B) Samples put in clamps with the side plates and, after tightening on four faces to firm pressure, they were returned to the same humidity as (A). All samples had the end grain exposed; two were set with the bolts pressing tangentially and two radially. This pressure slightly changed both the dimensions and the moisture contents, which were again measured.

(D)* Samples transferred, without altering the clamps, to a higher humidity (74·8 % R.H.).

(E) Side plates removed and samples returned to the same humidity as (D), thus allowing the samples to expand sideways and to absorb more moisture, the final restraint being unidirectional and, therefore, suitable for the calculation of Young's modulus. Two samples gave E_T and two E_R.

TABLE II.—SWELLING AND SORPTION OF SPRUCE CUBES

	Free	*Clamped*
cc./g. dry weight - - - - - -	V_1'	V_1''
moisture content - - - - - -	m'	m''

		A	B	D	E
R.H. -		32·8 %		74·8 %	
h (bar) -		0·00921		0·02098	
v (cc./g) -		116,300		51,000	
Clamps		Free	T and R	T and R	T
	Samples				
V_1'	1 and 3	2·115	2·118	2·203	2·203
	4 and 6	2·119	2·123	2·211	2·213
m'	1 and 3	0·0707	0·0707	0·1403	0·1409
	4 and 6	0·0700	0·0713	0·1402	0·1414
V_1''	2	2·130	2·112	2·146	2·167
	5	2·129	2·114	2·132	2·166
m''	2	0·0705	0·0684	0·1339	0·1357
	5	0·0701	0·0693	0·1333	0·1358

Four of the eight unstressed controls followed the clamped samples through to stage (E), the other four being left in the lower humidity (stage (B)). These were used for testing E_T and E_R mechanically. This test was done through the co-operation of the Timber Mechanics Section of the Laboratory.

It was hoped finally to release the samples completely from the clamps and to see whether they reverted to the moisture content and dimensions of the controls in the higher humidity, but this was not achieved owing to a breakdown. Also, for unexplained reasons, many of the measurements on radially clamped samples were found to be unsatisfactory and discordant, so that only the data on the tangentially clamped samples are available at the present stage. These are given in Table II.

From these experiments four groups of readings of sorption and dimensional changes can be obtained between the two humidities, i.e. A-D, B-D, A-E and B-E. In each case Δh is 42 % R.H. or $0{\cdot}0063 \times 10^6$ dynes/cm.^{2} at 23°c. Measurements on the controls gave s_1' and D' which, being constants of the material, should be approximately the same for all samples. From the clamped samples are obtained s_1'' and D'' whose values depend on an arbitrary degree of clamping and will therefore vary. Combining these free and clamped values in equation

---

* An intermediate humidity, stage (C), was also used, but it was found that the sorption interval was too small for accurate calculations.

(9) gives value of $(s_1)_h$ which should be constant. The results are given in Table III which shows that s_1' and D' are almost constant, but that $(s_1)_h$ shows a decided variation between the four columns. This is probably due to the variation of elastic constants with the magnitude of the stress applied or to plasticity because the larger values of $(s_1)_h$ tend to coincide with the smaller values of s_1'', i.e. with those samples where heavy stressing has produced the greatest diminution in sorption

Elastic Constants of the Gross Wood.—Before using this data to calculate the elastic constants of the wood substance, we may test the validity of the method by calculating the constants of the gross wood. In Table III will be found the values of k calculated from equation (4) and E_T from equation (5). The four pairs yield values of k which are seen to be reasonably concordant and to show a tendency for k to increase with s_1'', i.e. with decreasing stresses, thus supporting the suggestion made above in connection with the variation in $(s_1)_h$.

TABLE III.—HYGROSCOPIC AND ELASTIC CONSTANTS OF SPRUCE CUBES

R.H. 53·7 %
h 0·01509 (bar)
v 80,400 c.c./g.
m 0 1057 %
} mean values.

		Free	Clamped
Differential sorption	–	D'	D''
Differential swelling	–	s_1'	s_1''
do (at constant v.p.)	–	$(s_1)_h$	$(s_1)_h$
Bulk modulus (10^6 dynes/sq. cm.)	–	k_1'	—
Young's modulus („ „)	–	$E'T$	—

	A to D.	B to D.	A to E.	B to E
Samples 1 to 3				
s_1'	1·269	1·228	1·245	1·204
D'	0·1690	0·1691	0·1666	0·1666
s_1''	0·250	0·525	0·568	0·826
D''	0·1857	0·1798	0·1804	0·1749
$(s_1)_h$	11·652	12·349	9·984	8·801
k'_1	2344	2247	2946	3274
E'_T	by sorption		4981	
	by mech. test		4760	
Samples 4 to 6				
s_1'	1·309	1·275	1·251	1·217
D'	0·1676	0·1710	0·1648	0·1681
s_1''	0·045	0·282	0·557	0·778
D''	0·1863	0·1838	0·1791	0·1769
$(s_1)_h$	12·688	14·731	9·263	9·809
k_1'	2057	1823	2968	2897
E'_T	by sorption		6671	
	by mech. test		4160	

Only one pair of readings, viz. A and E, can be used to give a value of the Young's modulus E_T, as this is the only pair which involves a unidirectional stress. This is also given in Table III, together with the average values of E_T obtained by mechanical testing. The agreement is good for the first sample and poor for the second, though even for the latter it is of the right order of magnitude and the difference is not large when the variability of the material is taken into account.

These results can be taken as giving a satisfactory, if provisional, confirmation of the theory when applied to calculations made on the gross wood. Earlier experiments mentioned in the Report also support this conclusion.

Calculations of Poisson's ratio from the A and E stages were unsatisfactory as the side clamps used in the B and D stages appeared to have induced a permanent lateral distortion, to which the calculation of σ is, of course, very sensitive. The calculation of E is not so sensitive as σ to lateral restraint but it is possible that the high value obtained for E_T in samples 4 to 6, Table III, is due to this cause. The same cause probably explains the unsatisfactory results which are not reported here, obtained on the radially clamped samples, because here the wedges act tangentially, in which direction wood is most liable to suffer permanent distortion. This suggestion is supported by the fact that the A and E stages of the radially clamped samples both gave very high values of E_R. Experiments made without the use of the side plates are now in progress.

Elastic Constants of the Wood Substance.—In order to calculate the elastic constants of the stress-free wood substance we require values of V_3, s'_3, and $(s_3)_h$ at the mean R.H. of 53·8 %. These are obtained by interpolation from values given in Table I. The stress-free differential swelling s_3 is given by equation (15), the bulk modulus k_3 by equation (14) and the transverse rigidity modulus n_3 by equation (13). The transverse Young's modulus E_3 and the transverse Poisson's ratio σ_3 are calculated from k_3 and n_3 by equations (16) and (17). (See Table IV.)

TABLE IV.—HYGROSCOPIC AND ELASTIC CONSTANTS OF THE WOOD SUBSTANCE OF SPRUCE.

Stress-free data at 53·8 % R.H.

		A to D.	B to D.	A to E.	B to E.
Samples 1 to 3					
	s_3	0·937	0·936	0·943	0·938
	k_3	86500	86650	84750	85250
Transverse	n_3	2130	2530	2720	2105
	E_3	8320	9830	10,500	8210
	σ_3	0·952	0·943	0·939	0·951
Samples 4 to 6					
	s_3	0·934	0·934	0·937	0·941
	k_3	86100	87900	84400	85700
Transverse	n_3	1725	1705	1990	2430
	E_3	6760	6750	7780	9460
	σ_3	0·960	0·972	0·955	0·945

The longitudinal Young's modulus of the wood substance has, in the above calculations, been assumed to be very large compared to the transverse. In support of this, we may obtain an indication of its magnitude if we imagine the cell walls to be continuous hollow cylinders parallel with the grain. In this way the longitudinal Young's modulus $(E_L)_3$ can be obtained from the longitudinal modulus of the gross wood $(E_L)_1$ by putting $(E_L)_3 = (E_L)_1/(1 - A)$. Taking $(E_L)_1$ as the mean of five values of spruce at 12 % moisture content given by Carrington[5] i.e. $12·3 \times 10^{10}$ dynes per sq. cm., and A as 0·72, $(E_L)_3$ becomes $44·0 \times 10^{10}$ or more than 40 times the transverse E_3 above.

The stress-free elastic constants (Table IV) refer to a particular humidity which, for the natural wood, corresponds to a moisture content of about 10·6 % When, however, the stress free constants are calculated it must be remembered that these will no longer refer to the same moisture content, since the release of stress will also permit of greater sorption of water. By means of equation (4) we may calculate the stress-free values of D. This shows that the *rate* of sorption of the stress-free isothermal at this humidity is about double that for the natural material, but in order to obtain the new moisture content at this humidity we should require to know the values of D at all lower humidities and obtain the figure by summation. The present experiments give no indication of D at other humidities. This point will be discussed again in the next section.

---

[5] Carrington, *Phil. Mag.*, 1921, **41**, 848 , 1922, **43**, 871 ; 1923, **45**, 1055.

Variation with Moisture Content.—In common with other hygroscopic gels, the elastic constants of wood vary considerably with moisture content, the values of the Young's modulus and the rigidity modulus being reduced on increasing the moisture content. There is, therefore, some ambiguity in the above tables in taking mean values of the elastic constants from such a wide range of humidities. This was necessary in order to obtain measurable differences in the sorption and swelling of the clamped and free samples, but there is room for improvement in this direction, as at present it is not possible to obtain a range of mean humidities in order to study the variation of elastic constants with R.H.

At the time when the Report was written, no experimental determinations had been made of swelling under stress such as those just reported and, in order to give some idea of the elastic constants of the gel material, a method was evolved of calculating limiting values of the elastic constants. These limits were based on the limits to values of s_3 which, it was shown, must lie within the range of

$$s'_3 < s_3 < \frac{1 + n_3/k_3}{\text{[illegible]}}(s_3)_h. \qquad (18)$$

For k_3 the limits were fairly narrow, but for E_3 and n_3 the lower limit was zero (corresponding to a gel incapable of withstanding shear). On the evidence of similar calculations made from the present data, which are, however, confined to a single humidity, the value of k_3 was not much changed, but the upper limits of n_3 and E_3 were each about three times too high. In Table V these limiting values are reproduced from the Report, and serve to show the type of variation to be expected with moisture content.

TABLE V.—LIMITING VALUES OF THE DESORPTION MOISTURE CONTENT AND SOME ELASTIC CONSTANTS OF THE WOOD SUBSTANCE WHEN FREE FROM RESTRAINT. (TAKEN FROM "SWELLING STRESSES IN GELS.")

R.H. %	90	80	60	40	20	10
m (min.)*	0·201	0·177	0·138	0·102	0·065	0·040
(max.)	0·365	0·315	0·244	0·172	0·105	0·064
k_3 (min.)	12200	10600	9100	8500	8000	8000
(max.)	12300	10900	9600	9200	9800	10100
E_3 (min.)	0	0	0	0	0	0
(max.)	15400	16700	22800	31500	58000	83400
n_3 (min.)	0	0	0	0	0	0
(max.)	4000	4300	6000	9000	17600	28100

* These moisture contents appear low in comparison to those normally quoted for desorption because an attempt was here made to subtract the proportion of water held by capillary forces.

Part III. Discussion.

Strain Conversion.—Though the data here presented are far from complete, particularly in regard to the variation of properties with moisture content, two features stand out clearly. These are :

(1) that the bulk modulus of the gross wood k_1', Table III, is some 40 times smaller than that of the wood substance k_3, Table IV, and

(2) that the rigidity modulus of the wood substance n_3 is also very much smaller than k_3.

These two facts are related, because from the first equation in (11) we see that the bulk modulus of a hollow tube under external pressure is given by

$$\frac{1}{k_1} = \frac{1}{1 - A}\left\{\frac{1}{k_3} + \frac{A}{n_3}\right\}. \qquad (19)$$

From this it will be seen that if k_3 is much larger than n_3, the bulk modulus of the hollow tube under external pressure is almost entirely governed by the rigidity modulus of the wall material except in cases where A is very small. In the case of a wood cell we have also to take into account the reinforcement of the outer restraining sheath which, in fact, increases the bulk modulus of a natural cell by about four times, but since for spruce $(1 - A) = 0{\cdot}25$ roughly, the bulk modulus of the reinforced cell k_1' will turn out to be approximately equal to n_3. A comparison of Table III and Table IV shows this to be the case.

Thus natural wood has the property, unknown to any non-porous solid, of being able to absorb external hydrostatic stresses by means of shear strains in the wood substance. Since all types of stress in two-dimensional isotropic elasticity can be resolved into hydrostatic and shear components, it follows that wood must have the property of strain conversion for almost any type of stress that may be applied to it. It is generally accepted that pure hydrostatic strains are reversible, or elastic, and that any plastic losses occur under shear, so that the conversion of hydrostatic strains into shear strains will increase the probability of energy losses due to plastic deformation, and there can be little doubt that it is this property which gives wood an exceptionally high damping capacity to transverse vibrations.

Capillary Sorption.—This dependence of the strength of a tube on its rigidity modulus has an interesting and perhaps important application to a problem of capillary sorption in porous gels which emerged from a recent discussion with Dr. W. H. Banks of the Printing and Allied Trades Research Association. Since water, held at a reduced vapour pressure in a capillary, is under tension, it must transmit this tension to the walls of the tube. Imagine, therefore, a porous solid containing graded sizes of capillaries and initially saturated with water. If the humidity is gradually lowered, a tension is exerted on the capillary walls, which must tend to reduce their cross section. Thus, in a tube which would be expected to empty at a particular vapour pressure, evaporation will be delayed until a lower vapour pressure and a higher tension has been reached, and it is possible for the tube to reach a point of instability when the tension of the water is increasing faster than is the resistance of the tube to further compression.

Take, for mathematical simplicity, a system of circular capillary tubes of all sizes. If V_2 is the internal volume per cm. length of a particular tube and A is the cell space ratio (i.e. the proportion of void volume to total volume) it can be shown [1] that for an internal pressure p_2

$$\frac{dV_2}{V_2} = -\frac{dp_2}{(1 - A)}\left\{\frac{A}{k_3} + \frac{1}{n_3}\right\}, \quad . \quad . \quad . \quad (20)$$

where k_3 and n_3 are the bulk modulus and rigidity modulus of the material.

Since in wood substance, as in many other gels, n_3 is very small compared with k_3, the latter may be put equal to infinity in equation (20). If r_2 is the internal radius of the tube, $V_2 = \pi r_2^2$ and we have by integration, when $k_2 = \infty$.

$$A_{,} + \ln\left\{\frac{r_0^2}{r_2^2}(1 - A_0)\right\} = -\frac{p_2}{n_2}, \quad . \quad . \quad . \quad (21)$$

where r_0 and A_0 are the initial, unstressed, values of r_2 and A.

For capillaries which are large enough for the Kelvin equation to apply $-p = 2\tau/r_2$, τ being the surface tension and r_2 the final equilibrium value of the internal radius.

Fig. 4 gives the graph of r_0 against r_2 for two materials, in each of which A_0 has arbitrarily been fixed at $0{\cdot}5$. In the first, $n_A = 6000 \times 10^6$ dynes/sq. cm. and corresponds to the rigidity of wood substance at a low moisture content. In the second, $n_B = 15 \times 10^6$ and approximates to the rigidity of soft rubber. The graph shows the minimum values of r_0 to be about 10^{-8} cm. at 1 % R.H. for A; and about 5×10^{-6} at 98 % R.H. for B.

These are the smallest initial, i.e. unstressed, radii which are capable of being dried under a falling humidity without being reduced to zero cross section by the ever-increasing tension of the water. Below the critical R.H. no capillary water can be present and no further capillary effects can be manifested since the larger capillaries are empty and the smaller ones have been eliminated by collapse.

These curves cannot, of course, be applied quantitatively to real materials where the capillary structure is irregular, nor to such small and meaningless values of the capillary radius as that obtained for curve A above. The interest lies more in the qualitative significance of curve B, which shows that voids in soft materials can collapse at such high humidities that the corresponding pore structure should be visible in the microscope. At these high humidities, criticisms about the reality of surface tension effects in small capillaries, which are rightly levelled against the blind application of the Kelvin equation, can hardly present any difficulty.

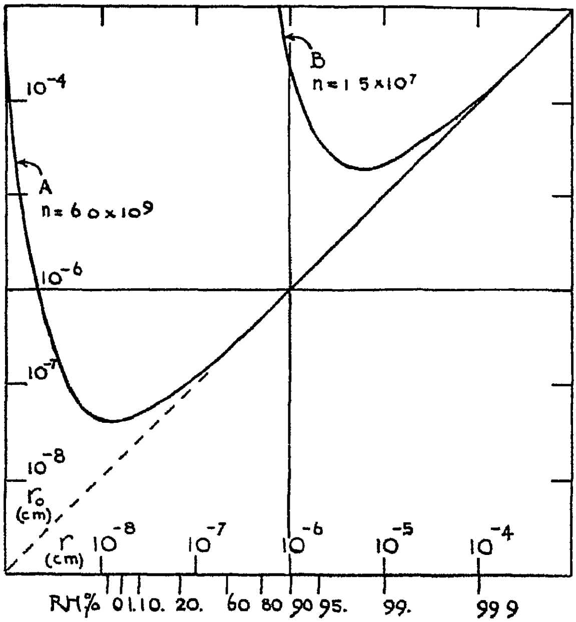


FIG. 4.

The bulk modulus, even of soft gel substances, does not fall below that of water, i.e. $20{,}000 \times 10^6$ dynes per sq. cm., and thus will be large compared to the rigidity of any of the softer gels. It is evident, therefore, that soft gels of very low rigidity may suffer capillary collapse at humidities even closer to saturation than curve B above and this will produce abnormally large shrinkages at high moisture contents.

In seasoning wet wood at elevated temperatures there is a tendency for anomalous shrinkage to occur in the early stages of drying and microscopic examination reveals a considerable distortion in the cell cavities, a condition which is known as collapse. In wood at elevated temperatures the rigidity of the cell wall near saturation will be much lower than that reported here for a low humidity at room temperature, and may well approximate to that of the rubber. Thus it seems likely that it is the low rigidity which leads to the collapsed condition, not necessarily due entirely to the tension in the comparatively large capillaries which can be observed, but to the general distortion following the elimination of all the smaller capillaries in the material. Previous explanations of collapse in timber have been attempted on the idea that the walls of an anatomical cell, completely filled with water, will collapse if no air is able to enter when the water evaporates, but this does not satisfactorily account for collapse in cases where the cells may not be completely waterlogged. It now appears that the exclusion of air is not necessary to the explanation.

Sorption Hysteresis.—In the list of distinguishing properties of gels it was pointed out that gels differed from true solutions in possessing rigidity and in showing sorption hysteresis. Since plastic hysteresis is a manifestation of work loss in a cycle of shear stresses, and sorption hysteresis of work loss in a cycle of adsorption and desorption, it seems probable that an explanation of the presence of the latter in gels and of its absence in solutions should be sought in terms of plasticity in rigid gels.

The fact that wood substance does not disperse into solution at the saturated vapour pressure, although the material has hygroscopic affinities similar to that of glucose, means that somewhere in the structure of the system there are permanent cross-linkages or entanglements between the long chain molecules, of sufficient strength to prevent dispersion. We may therefore look on solution as the only "stress-free" state of swelling.

The fundamental equation connecting the change of differential sorption, from D in the free state to D' when under stress, was given above (equation (4)).

$$D' - D = \frac{ks_p}{Vv}(s_p - s'), \quad . \quad . \quad . \quad . \quad (22)$$

where s_p and s' refer to the corresponding free and stressed values of the differential swelling. Since all quantities here are constant except D' and s', a change in D' due to hysteresis presupposes a change in s' between adsorption and desorption.

It has also been described in Part I how s' is determined by the mechanical properties of the gel and of the type of stress imposed on it. For example, in the particular case of an isotropic gel enclosed in a cylinder, the movement of a frictionless piston per g. of water adsorbed was given by [1]

$$s'_x = \frac{3k}{3k + 4n}s_p. \quad . \quad . \quad . \quad . \quad (23)$$

Suppose now that the gel is plastic, then the rigidity of the gel, which is the slope of the stress-strain curve in shear, will appear to be different in extension and compression. If the load on the piston increases on swelling then the gel becomes more compressed on adsorption and more extended on desorption, so that we shall obtain different values of n for use in the above expression on adsorption and desorption and hence different values of s'_x. This will lead to a hysteresis loop in D'.

This particular illustration is sufficient to show that any type of swelling constraint involving shear strains in a plastic gel will produce hysteresis in the sorption isothermal. From the fact that no hysteresis occurs in true solutions, where no shear strains are possible, we may infer that plasticity is the cause of sorption hysteresis in gels.

Hysteresis at a Saturated Vapour Pressure.—Plasticity also explains why the hysteresis loop does not necessrrily close at the saturated vapour pressure, as in Fig. 1. For a true solution $D = 0$ and $m = \infty$ at the saturated vapour pressure, so equation (22) shows that, for a solution (or a gel) under stress, D' is positive and m is finite at saturation. Thus the sorption curve must cut the saturated vapour pressure line at a finite angle and cannot run to infinite dilution unless the saturated vapour pressure is exceeded. This, in normal circumstances, is impossible. Thus, for a gel under stress, the limit of sorption from vapour is set by the properties of the vapour, which cannot exert a pressure in excess of its saturated value. For a plastic gel under stress, therefore, the saturated vapour pressure line cuts the hysteresis loop at *two* points and the adsorption and desorption curves are not coincident (Fig. 5). In order to reach the waterlogged state from the adsorption side, it is necessary that the vapour pressure should be temporarily raised above the saturation value. This can be done by forcibly immersing the gel and inducing momentary high liquid pressures by agitation. In support of this suggestion, the following experiment may be quoted. A small chip of wood, previously vacuum-dried, is held over water in an air-free enclosure, until it has reached saturation equilibrium at about 35 % moisture content. It is then dropped on to the water surface, when it will be found to float high on the surface like a " water boatman " without wetting. If it is now violently shaken it will suddenly become waterlogged and sink. After this treatment it will hold as much as 100 % moisture content even after the vapour pressure has been slightly reduced below saturation.

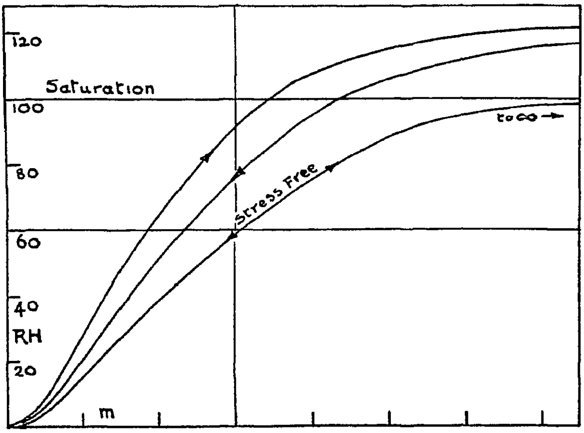


FIG. 5.

Summary.

The physical property which chiefly distinguishes gels from true solutions is that they possess rigidity and can therefore resist directional stresses.

Exact expressions are given connecting the change of sorption and of anisotropic swelling in a gel with the change of stresses applied to it. Equations are derived by which the data on sorption and swelling under restraint may be used to calculate the elastic constants of a gel possessing rhombic symmetry. The theory is equally applicable to any other system of elastic symmetry.

No stresses need be measured directly as the stress and strain are functions respectively of the reductions in sorption and in swelling.

The theory is applied to measurements made on the swelling of spruce wood in specially designed clamps and, from these, calculations are made of the principal elastic constants both of the natural wood and of the gel material composing the cell wall.

The plasticity of gels under shear stress provides a satisfactory explanation of the occurrence of sorption hysteresis in gels and of its absence in true solutions. It is shown that soft gels containing capillary spaces may reach a point of instability on drying, such that all voids below a critical size will collapse when the falling humidity reaches a critical value.

Physics Section,
Forest Products Research Laboratory,
Princes Risborough, Bucks.

SWELLING STRESSES IN ANISOTROPIC CYLINDERS

By F. L. Warburton.

Received 18th September, 1946.

1. Introduction.

Many natural fibres are hygroscopic gels of roughly cylindrical form, and their principal longitudinal surfaces are a series of concentric cylinders and a series of radial planes. When such a gel absorbs water or other liquid, the swellings in the planes and surfaces in directions perpendicular to the axis are geometrically constrained to be the same, and if the free swellings in the radial and tangential directions are different, a system of internal stresses will be set up to give a quasi-equilibrium. Plastic flow may ultimately relieve the stresses and give a final equilibrium sorption, but the approach to the final equilibrium will depend on the stresses and, therefore, on the geometrical form of the gel.

The following analysis deals with this problem for small changes in the amount of absorbed liquid when it is possible to apply the elastic theory for small strains. A complete investigation of the stresses set up for large changes will involve the theory of finite strains, and requires separate analysis.

2. Derivation of Stresses.

The stress-free swelling of an anisotropic gel in the form of a symmetrical cylinder gives distortions which are inconsistent with the condition of compatibility and therefore additional distortions must be produced by the existence of internal stresses. These distortions will be of two kinds, namely elastic distortions and distortion due to the fact that in order to maintain thermodynamic equilibrium the amount of absorbed substance will depend on the stresses. In order to simplify the calculations it will be assumed that over the range of contents of absorbate concerned, the variation of equilibrium vapour pressure with amount of absorbate is constant, and that the elastic distortions are sufficiently small for ordinary elastic theory to apply. These assumptions will be approximately correct except, near the centre of the cylinder, for large changes in absorbate content. It will also be assumed that the swelling takes place from a stress-free state.

It will first be necessary to obtain an expression for the change in moisture content with applied stresses at constant temperature and constant vapour pressure. The necessary equation has already been obtained by Barkas [1] and the author [2] for cases in which the initial shear stresses are small compared with the rigidity modules. With appropriate change of notation this can be written

$$V(s_r P_r + s_\theta P_\theta + s_z P_z) = v(\partial p'/\partial n)\delta n - v\delta p'. \quad . \quad . \quad (1)$$

where V is the volume of the gel containing n moles of absorbate, s_r, s_θ and s_z are the fractional swellings per mole in the radial, tangential, and longitudinal directions respectively, P_r, P_θ and P_z are the stresses in these directions (tensions being positive); v is the molar volume of the absorbate in the vapour state and p' is the equilibrium vapour pressure.

---

[1] Barkas, *Swelling Stresses in Gels, F.P.R. Special Report No.* 6, H.M.S.O. 1945.

[2] Warburton, *Proc. Physic. Soc.*, 1946, **58**, 585.

The signs on the L.H.S. are reversed, because the thermodynamic convention was used for stresses in the original equation, i.e. with pressures positive. For gels of the type considered, s_zP_z can be neglected and, therefore, for variations at constant vapour pressure.

$$v(\partial p'/\partial n)\delta n = V(s_rP_r + s_\theta P_\theta) \quad . \quad . \quad . \quad (1.1)$$

$(\partial p'/\partial n)$ is the change in vapour pressure with amount of absorbate at constant stress and temperature; i.e. it is proportional to the slope of the isotherm, which can be assumed a straight line provided the moisture content is neither very small nor approaches saturation. (The product $v(\partial p'/\partial n)$ is proportional to the absolute temperature.)

Hence
$$\delta n = \frac{V}{v(\partial p'/\partial n)}(s_rP_r + s_\theta P_\theta)$$
$$= \beta(s_rP_r + s_\theta P_\theta) \quad . \quad . \quad . \quad . \quad . \quad . \quad (1.2)$$

where
$$\beta = V/v(\partial p'/\partial n) \quad . \quad . \quad . \quad . \quad . \quad . \quad (1.3)$$

When a cylinder which is initially stress-free absorbs sufficient absorbate, $(\Delta n + \delta n)$, to maintain it in equilibrium with a rectangular slab which increases its absorbate content by Δn moles for the same quantity of gel, the extensions of a small portion of the cylinder after equilibrium has been attained will be ϵ_r, ϵ_θ, and ϵ_z respectively, neglecting the longitudinal swelling s_z where,

$$\begin{aligned}
\epsilon_r &= s_r(\Delta n + \delta n) + P_r/E_r - \sigma_{r\theta}P_\theta/E_\theta - \sigma_{rz}P_z/E_z, \\
\epsilon_\theta &= s_\theta(\Delta n + \delta n) + P_\theta/E_\theta - \sigma_{\theta r}P_r/E_r - \sigma_{\theta z}P_z/E_z, \\
\epsilon_z &= P_z/E_z - \sigma_{zr}P_r/E_r - \sigma_{z\theta}P_\theta/E_\theta \quad . \quad . \quad . \quad . \quad (2)
\end{aligned}$$

where E_r, E_θ, and E_z are the values of Young's modulus in the three directions, and $\sigma_{r\theta}$ is the ratio of the contraction in the r direction to extension in the θ direction for stresses in the θ direction, etc. Eliminating P_z between equations (2) and substituting for δn from equation (1.2), gives:

$$\begin{aligned}
\epsilon_r &= s_r\{\Delta n + \beta(s_rP_r + s_\theta P_\theta)\} + (1 - \sigma_{rz}\sigma_{zr})P_r/E_r - (\sigma_{r\theta} + \sigma_{rz}\sigma_{z\theta})P_\theta/E_\theta - \sigma_{rz}\epsilon_z \\
\epsilon_\theta &= s_\theta\{\Delta n + \beta(s_rP_r + s_\theta P_\theta)\} - (\sigma_{\theta r} + \sigma_{\theta z}\sigma_{zr})P_r/E_r + (1 - \sigma_{\theta z}\sigma_{z\theta})P_\theta/E_\theta - \sigma_{\theta z}\epsilon_z \quad (3)
\end{aligned}$$

The condition of compatibility is:

$$\epsilon_r = \epsilon_\theta + r\frac{d\epsilon_\theta}{dr} = d(r\epsilon_\theta)/dr \quad . \quad . \quad . \quad . \quad (4)$$

Eliminating ϵ_r and ϵ_θ from equations (3) by means of relation (4) and using the condition that ϵ_z is independent of r, as it must be for a cylinder except near the ends, where P_z is zero:

$$\begin{aligned}
&\{\beta s_rs_\theta - (\sigma_{\theta r} + \sigma_{\theta z}\sigma_{zr})/E_r\}\frac{d(rP_r)}{dr} + \{\beta s_\theta^2 + (1 - \sigma_{\theta z}\sigma_{z\theta})/E_\theta\}\frac{d(rP_\theta)}{dr} \\
&\quad - \{\beta s_r^2 + (1 - \sigma_{rz}\sigma_{zr})/E_r\}P_r - \{\beta s_rs_\theta - (\sigma_{r\theta} + \sigma_{rz}\sigma_{z\theta})/E_\theta\}P_\theta \\
&\quad = \Delta n(s_r - s_\theta) - \epsilon_z(\sigma_{rz} - \sigma_{\theta z}) \quad . \quad . \quad . \quad (5)
\end{aligned}$$

The equation of static equilibrium is[3]:

$$P_\theta = \frac{d(rP_r)}{dr} \quad . \quad . \quad . \quad . \quad (6)$$

Eliminating P_θ between (5) and (6) and multiply by r

$$\begin{aligned}
&\{\beta s_\theta^2 + (1 - \sigma_{\theta z}\sigma_{z\theta})/E_\theta\}(rd/dr)^2(rP_r) + \{\beta s_rs_\theta - (\sigma_{\theta r} + \sigma_{\theta z}\sigma_{zr})/E_r - \beta s_rs_\theta \\
&\quad + (\sigma_{r\theta} + \sigma_{rz}\sigma_{z\theta})/E_\theta\}(rd/dr)(rP_r) - \{\beta s_{r_2} + (1 - \sigma_{rz}\sigma_{zr})/E_r\}(rP_r) \\
&\quad = (s_r - s_\theta)r\Delta n - (\sigma_{rz} - \sigma_{\theta z}) + \epsilon_z.
\end{aligned}$$

---

[3] Timoshenko, *Theory of Elasticity* (McGraw-Hill, New York, 1934), p. 367.

But $\sigma_{\theta r}/E_r = \sigma_{r\theta}/E_\theta$; $\sigma_{z\theta}/E_\theta = \sigma_{\theta z}/E_z$; and $\sigma_{zr}/E_r = \sigma_{rz}/E_z$, and the coefficient of the first differential vanishes, hence writing A_θ for

$$\{\beta s_\theta^2 + (1 - \sigma_{\theta z}\sigma_{z\theta})/E_\theta\}$$

and similarly for A_r, one obtains

$$A_\theta\left(r\frac{\mathrm{d}}{\mathrm{d}r}\right)^2(rP_r) - A_r(rP_r) = (s_r - s_\theta)r\Delta n - (\sigma_{rz} - \sigma_{\theta z})r\epsilon_z \quad . \quad (7)$$

This is a homogeneous linear equation of which the solution is

$$rP_r = B_1 r^\alpha + B_2 r^{-\alpha} + cr, \text{ where } \alpha = \sqrt{A_r/A_\theta}$$

c is obtained by differentiation of the R.H.S. of equation (7) and B_1 and B_2 can be found from the end conditions. For a solid cylinder these are B_2 is zero since P_r must be finite when $r = 0$, and $B_1 a^{\alpha-1} = -c$ if $P_r = 0$ at the outside where $r = a$,

i.e.
$$P_r = c\{1 - (r/a)^{\alpha-1}\}$$

where c as found by differentiation is $\{(s_r - s_\theta)\Delta n - (\sigma_{rz} - \sigma_{\theta z})\epsilon_z\} \times (A_\theta - A_r)$. Hence the complete solution is:

$$\begin{aligned} P_r &= c\{1 - (r/a)^{\alpha-1}\} \\ P_\theta &= c\{1 - \alpha(r/a)^{\alpha-1}\} \\ P_z &= E_z\epsilon_z + c(\sigma_{rz} + \sigma_{\theta z}) - c(r/a)^{\alpha-1}(\sigma_{rz} + \alpha\sigma_{\theta z}) \quad . \quad . \quad (8) \end{aligned}$$

ϵ_z can be obtained from these equations by the fact that the resultant stress over the circular cross-section is zero, *i.e.* $\int_0^a 2\pi r P_z \mathrm{d}r = 0$.

However, if $\sigma_{rz} = \sigma_{\theta z}$, ϵ_z does not enter into the equations for P_r and P_θ.

3. Effect of Internal Stresses on the Isotherm.

From the expressions for P_r and P_θ it is possible to calculate the mean value of δn for the cylinder; i.e. the effect of the stresses on the isotherm.

The mean value for axial symmetry is:

$$\overline{\delta n} = \int_0^a (2\pi r\delta n)\mathrm{d}r/\pi a^2 = (2/a^2)\int_0^a (r\delta n)\mathrm{d}r.$$

Substituting for δn in terms of P_r and P_θ from equation (1.2) and for P_r and P_θ from equations (8):

$$\overline{\delta n} = (2/a^2)\int_0^a \beta c\left\{s_r\left(r - \frac{r^\alpha}{a^{\alpha-1}}\right) + s_\theta\left(r - \alpha\frac{r^\alpha}{a^{\alpha-1}}\right)\right\}\mathrm{d}r$$

$$\overline{\delta n} = \beta c\frac{\alpha - 1}{\alpha + 1}(s_r - s_\theta) \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

4. Approximate Values for Water Absorption by Keratin.

Experimental work on this subject is as yet incomplete but sufficient data are available to obtain an approximate estimate of the stress set up for small changes in the water sorption of keratin cylinders, and thereby to obtain the order of the effects likely to occur. For this purpose it will suffice to neglect the term $(\sigma_{rz} - \sigma_{\theta z})\epsilon_z$ in comparison with $(s_r - s_\theta)\Delta n$.

The percentage moisture regain gives the number of moles of water absorbed by 1800 g. of the dry material, and the calculations will refer to this amount of dry keratin. V in equation (1.3) is then the volume containing 1800 g. of dry keratin at the regain before the sorption change occurs. Taking the initial conditions as 25° C. and 70 % R.H., the regain is 15 %, and as the density of keratin[4] under these conditions is 1·313 g./cc., V is 1578 cc. v, the molar volume of water vapour at 25° C. and

---

[4] King, *J. Text. Inst.*, 1926, **17**, T53.

70 % R.H. is 1.12×10^6 cc., and the gradient of the isotherm, $\partial p'/\partial n$, is 1900 dynes/cm.^{2} per mole. β is, therefore, approximately 7×10^{-7} in c.g.s. units. The free tangential swelling, s_θ, as measured for horn keratin is roughly 2×10^{-3} and the free radial swelling is 5×10^{-3} per mole absorbed by 1800 g. of dry keratin.[5] E_θ and E_r are[5] approximately 4×10^{10} dynes/cm.^{2}, and $\sigma_{\theta z}$ is approximately 0·25. The same value can be used for the other Poisson's ratios without appreciably altering A_θ and A_r. Using these values, one obtains: $A_\theta \simeq 28 \times 10^{-12}$, $A_r \simeq 43 \times 10^{-12}$, $\alpha \simeq 1{\cdot}25$, and $c \simeq -2 \times 10^8 \Delta n$.

Equations (8) then give:

$$P_r = -2 \times 10^{-8}[1 - (r/a)^{0.25}]\Delta n$$
$$P_\theta = -2 \times 10^8[1 - 1{\cdot}25(r/a)^{0.25})\Delta n.$$

The hydrostatic pressure at the centre of the cylinder due to unit change of regain is, therefore, 200 bar, and there is a tangential tension at the surface of 50 bar.

The change in the slope of the isotherm due to these stresses is given by equation (9); i.e.

$$\overline{pn} = \left\{7 \times 10^{-7} \times (-2 \times 10^8) \times \frac{0{\cdot}25}{2{\cdot}25} \times 3 \times 10^{-3}\right\}\Delta n = -4{\cdot}5 \times 10^{-2}\Delta n \quad . \quad (10)$$

The change in absorption at the centre due to the mechanical forces is given by equation (1.2) as $\beta(s_r P_r + s_\theta P_\theta)$, and numerical values give this coefficient as roughly (-1). There is, therefore, very little change in regain at the centre of the cylinder, a quasi-equilibrium being attained mainly by the imposition of mechanical forces. Plastic flow will ultimately relieve the mechanical stresses to give the final equilibrium regain throughout the cylinder, but equation (10) shows that the final average regain will be closely attained for cylinders of keratin before the release of mechanical stresses occurs. A cylinder of keratin should thus show the bulk of its regain change rapidly, because the stress system gives an excess absorption in the outer layers and a deficiency near the centre where diffusion slows up the sorption process, and the slow levelling-out process will, according to the above calculations, give little change in the total sorption.

These conclusions appeared to be borne out by a wool fibre absorbing water vapour: changed humidity conditions give a rapid change in regain, and this is followed by a very slow approach to final equilibrium and constant weight.

Summary.

The mechanical stresses set up in cylinders of hygroscopic gels due to the anisotropy of transverse swelling are considered for small changes in the amount of absorbed liquid. It is shown that these result in a small change in the slope of the isotherm, and that a quasi-equilibrium is attained in the centre by mechanical forces rather than by absorption of liquid. The final equilibrium with a uniform distribution of liquid through the cylinder is obtained when plastic flow releases the swelling stresses, and the change from quasi- to final equilibrium will be much slower than the establishment of the quasi-equilibrium.

I am indebted to Mr. B. H. Wilsdon, Director of Research, for encouragement in this work, to Dr. A. B. D. Cassie for discussion, and to the Council of the Wool Industries Research Association for permission to publish this account.

Wool Industries Research Association,
Torridon,
Headingley, Leeds 6.

[5] Measurements by King and Warburton in W.I.R.A. Laboratories.

THE STRUCTURE AND THE MECHANISM OF DEFORMATION IN CELLULOSE GELS.

I. GENERAL PICTURE OF A MOLECULAR NETWORK.

By P. H. Hermans and D. Vermaas.*

Received 13th August, 1946.

1. Introduction.

Taking the point of view that gels from flexible linear macromolecules consist essentially of molecular network structures with fixed junction points (cross links either of a chemical or of a physical nature), there are a number of consequences regarding the nature of the structural changes upon swelling, shrinking and deformation which have recently been recognised in the field of vulcanised rubber-like polymers.[1–7] Both swelling and deformation are associated with changes in the spatial arrangement of the junction points and the degree of kinkiness of the molecular chains interconnecting them. This picture has led to a rather successful quantitative theoretical treatment on the basis of thermodynamics and the statistical configuration of chain molecules combined with certain geometrical and mechanical assumptions.

In our own investigations in war-time on the behaviour of cellulose gels we have independently been led to a quite analogous picture and in this paper some characteristic features of this work will be offered in order to show that the general ideas recently developed in rubber-like polymers may, to a further extent than heretofore recognised, be transferred to the physics of macromolecular gels in general.

True, the quantitative treatment presents greater difficulties in the case of cellulose for obvious reasons (much greater intermolecular forces and lack of reversibility of deformation). On the other hand, deformation in the swollen state plays a predominant part and can easily be studied here, in contrast to rubber where these conditions are more difficult to realise and have hitherto almost entirely been neglected. Furthermore, in the case of rubber, the cross-linking giving rise to the network formation usually takes place in the " dry " polymer. It is then reasonable to assume that in this " dry " condition the degree of kinkiness of the chains corresponds to the " normal " one given by Kuhn's distribution function and may be designated as " normal coiling." Swelling is then necessarily associated with a certain amount of " uncoiling ".

In cellulose gels cross-linking occurs in solution at the sol-gel transformation, i.e. at the formation of the "*primary gel*" and the degree of swelling q_0 of the primary gel (this being the ratio between the volume of the gel and that of the dry polymer) can be varied at will to a considerable extent by varying the concentration of the sol. Shrinking (as e.g. occurs in the drying of a swollen isotropic gel) will then be bound up with a "supercoiling" and the configuration of the chains in the shrunken "secondary" gel will never be "normal".

---

* Communication N2.40 from Lab. for Cellulose Research, A.K.U. and Affiliated Companies, Utrecht, Holland.

[1] Busse, *J. Physic. Chem.*, 1932, 36, 2862.

[2] Kuhn, *et al.*, *Kolloid. Z.*, 1934, 68, 2 ; 1942, 101, 248 ; *J. Polymer Sci.* (in press).

[3] Wall, *J. Chem. Physics*, 1942, 10, 485.

[4] Flory and Rehner, *ibid.*, 1943, 11, 521.

[5] Treloar, *Trans. Faraday Soc.*, 1943, 39, 36, 242 ; 1944, 40, 109 ; 1946, 42, 83.

[6] J. J. Hermans, *Kolloid. Z.*, 1943, 103, 210.

[7] J. J. Hermans, *Trans. Faraday Soc.*, (in preparation).

Assuming that the coiling in all primary gels prepared in a similar fashion is the same (though perhaps not necessarily the "normal" one), it will be clear that the structure of "secondary" gels at a given degree of swelling q and their behaviour on deformation must be essentially dependent on the degree of swelling of the primary gel.[8]

2. Influence of the Primary Degree of Swelling and of Volume Changes on the Anisotropy upon Deformation.

Isotropic swelling and shrinking can be considered as an isotropic spatial expansion or compression of the junction-point pattern and deformation as an anisotropic expansion giving rise to configurational optical anisotropy. In order to see in what way volume changes interfere with anisotropy changes upon unidirectional deformation, we shall first consider a molecular network, following Kuhn[3] *et al.* Their theory, which was confined to constant volume, leads to the following formula for the anisotropy of polarisability at a relative unidirectional elongation v of the primary isotropic gel:

$$\beta_{\parallel} - \beta_{\perp} = k(v^2 - v^{-1}) \quad . \quad . \quad . \quad . \quad (1)$$

where k is a constant characteristic for the gel. This theory can be readily modified to allow for changes in degree of swelling of the primary gel. We then find:

$$(\beta_{\parallel} - \beta_{\perp})_f = k(q_i/q_0)^{2/3}\,(v^2 - v^{-1} \,.\, q_f/q_i) \quad . \quad . \quad . \quad (2)$$

where q_i and q_f are the initial and final degree of swelling before and after deformation,[9] while q_0 is the degree of swelling where coiling is normal. The equation also applies to isotropic swelling or shrinking. Then v (the rel. change in linear dimensions) is, of course, $(q_f/q_i)^{1/3}$. Substituting this in (2), we find $\beta_{\parallel} - \beta_{\perp} = 0$; the body remains isotropic.

Designating $\beta_{\parallel} - \beta_{\perp}$ as the "orientation" of the network, equation (2) shows that the orientation reached at a given elongation v depends on the volume change q_f/q_i occurring on stretching and, at a given q_i and volume change the orientation will be inversely proportional to the 2/3power of q_0.

3. Application to Isotropic Cellulose Filaments.

In the case of highly swollen isotropic cellulose model filaments[10] there is always a considerable decrease in volume upon stretching.[11] The change in length and volume of such filaments on isotropic shrinking and on deformation are diagrammatically shown in Fig. 1. After extension from the initial length l_i to the length l_f, the conventional rel. elongation $v = l_f/l_i$. Since the volume drops from q_i to q_f and this change in swelling as such involves a change in linear dimensions, the final length l_f of the stretched filament should be compared with the length $(l_{iso})_f$ of the filament isotropically shrunken to q_f, i.e. with $(q_f/q_i)^{1/3}l_i$.

---

[8] Various striking illustrations of the "memory" of gels for the degree of swelling of the primary gel, which can be understood in this way, were already reported in early literature, e.g. Gortner and Hoffmann, *J. Amer. Chem. Soc.*, 1921, **43**, 2199; *Proc. Soc. Exp. Biol. Med.*, 1922, **11**, 257; *J. Physic. Chem.* 1927, **31**, 464.

[9] The derivation of this equation, which involves no difficulties, is indicated in ref. 7.

[10] P. H. Hermans, and de Leeuw, *Kolloid Z.*, 1937, **81**, 321.

[11] P. H. Hermans, *Cellulosechemie*, 1941, **19**, 122.

A more rational degree of extension, hence, is given by

$$v_a = l_f/(l_{iso})_f = v(q_i/q_f)^{1/3}. \quad . \quad . \quad . \quad . \quad (3)$$

(This is the same measure of elongation as introduced by Kratky *et al.*[12] in his theory of affine deformation.) Substituting in (2) we get

$$(\beta_{\parallel} - \beta_{\perp})_f = k(q_f/q_0)^{2/3}(v_a^2 - v_a^{-1}). \quad . \quad . \quad . \quad (4)$$

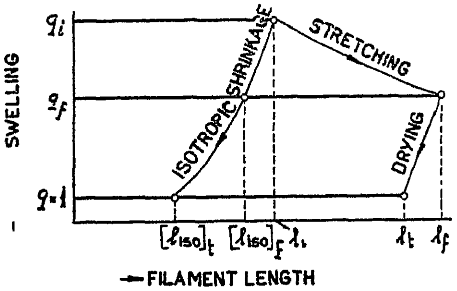


FIG. 1 —Diagram showing the relation between degree of swelling q and length of swollen isotropic filaments on isotropic shrinking and on stretching followed by drying.

Measurements of double refraction are, however, carried out in the air-dry state. Now, if the filament is dried to a degree of swelling $q = 1$, its length contracts from l_f to $l_t = \mu l_f$ cf. Fig. 1), where μ is the relative longitudinal contraction. The rational measure of elongation in the dry state will then be $v_t = l_t/(l_{iso})_t = \mu q_i^{1/3}\ l_f/l_i$, since $(l_{iso})_t$ $q_i^{-1/3}l_i$. On the other hand, the elongation of the dry fibre with respect to the initial swollen state (q_i) is $l_t/l_i = v_t q_i^{-1/3}$. If we apply the equation (2) to the dry state, we obtain therefore for the anisotropy in this state (since now $q_f = 1$):

$$(\beta_{\parallel} - \beta_{\perp})_t = kq_0^{-2/3}(v_t^2 - v_t^{-1}). \quad . \quad . \quad . \quad (5)$$

This implies that the anisotropic shrinking from $q = q_f$ to $q = 1$ is considered as a deformation of essentially the same character as that brought about by stretching. In fact, we assume that the relation (2) is of general applicability, irrespective of whether the change in the linear dimensions is caused by stretching or by drying. For small elongations eq. (5) may be written

$$(\beta_{\parallel} - \beta_{\perp})_t = 3kq_0^{-2/3}(v_t - 1). \quad . \quad . \quad . \quad . \quad (6)$$

showing that the initial slope of the anisotropy against v_t curves should be inversely proportional to the 2/3 power of the degree of swelling q_0 where coiling is normal, which presumably is the degree of swelling of the primary gel from which the filaments are prepared.

The theory does not specify the change in orientation upon drying. This is determined by the magnitude of the contraction factor μ. If the orientation remains constant on drying this means that

$$(\beta_{\parallel} - \beta_{\perp})_f = (\beta_{\parallel} - \beta_{\perp})_t.$$

It is obvious that this amounts to an equation for μ.

4. Experimental.

A series of isotropic xanthate filaments[10] was prepared from nine viscoses of different composition. The concentration of free alkali and of "salt"[13] as well as the "xanthate ratio"[14] and the coagulating bath (an ammonium sulphate solution of the density ~ 1·08) were kept constant, whereas the cellulose concentration and the degree of polymerisation (DP) were varied as follows :

---

[12] Baule, Kratky and Treer, *Z. physik. Chem. B.*, 1941, 50, 255.

[13] The "salt" in viscose is a mixture of Na_2CO_3 and Na_2CS_3 in the molar ratio 1 : 2, which is formed when CS_2 acts on NaOH.

[14] The number of xanthate groups per glucose unit, which was always near to 0·37 at spinning.

Viscose no.	1	2	3	4	5	6	7	8	9
% cellulose	2	4	4	6	6	6	8	8	10
DP	660	270	390	200	270	390	200	270	170
q_0	29·0	16·1	16·0	13·4	13·1	10·5	10·5	9·8	8·1

The observed degree of swelling q_0 of the primary xanthate filaments is given in last line of the Table. The filaments were stretched to various extents, then transformed into cellulose and dried, accurately measuring their length and volume in the initial dry, and intermediate states. The birefringence of the filaments was measured in the dry state and plotted against the "dry elongation" v_t. The result is given in Fig. 2, every viscose yielding a linear relation of the form as given in eqn. (6), the slopes of the curves lying in the same order as the values of q_0.

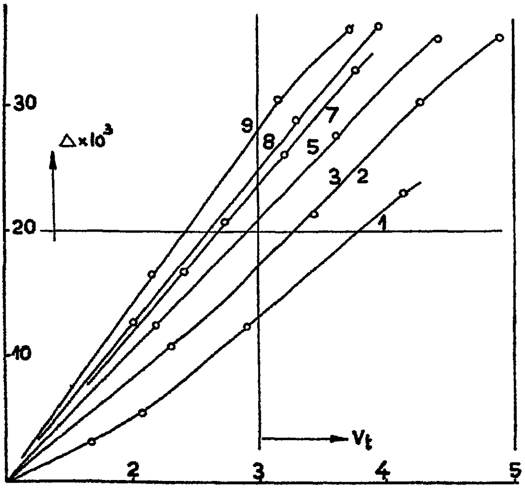


FIG. 2.—Double refraction Δ of the dry filaments plotted against the elongation v_t. For the sake of clarity some curves have been omitted ; they fit in well with the other curves.

According to eqn. (6), the initial slope should be proportional to $q_0^{-2/3}$. A check is given in the logarithmic plot of Fig. 3. The slope of this plot, i.e. the exponent of q_0, cannot be determined with great accuracy, but can be shown to lie between 0·6 and 0·9.[16] We can say, therefore, that the antiparallelism between the slope and q_0 is borne out by experiment ; whether this slope is strictly proportional to $q_0^{-2/3}$ cannot be decided with any certainty. A similar result was obtained when the fibres were first decomposed to cellulose and then stretched. The product of the slope and the value of $q_0^{-2/3}$ is somewhat larger in this case, and the values of this product scatter more strongly about the mean.

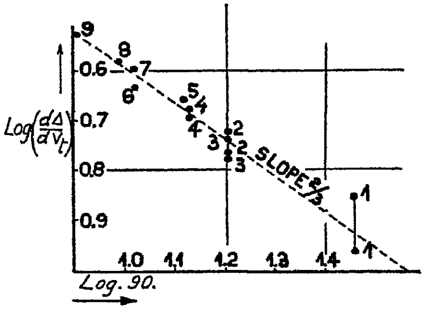


FIG. 3.—Plot of the logarithm of the slope of the curves in Fig. 2 against the logarithm of the swelling degree q_2 of the primary gel.

From X-ray measurements which cannot be communicated here, evidence is obtained that the orientation is not changed materially upon drying. Calculating the factor μ of longitudinal contraction on this basis from (4) and (5) we obtained values which tallied with the experimental ones within 10 %.

---

[16] The gel from viscose No. 1 was very weak and no accurate measurements could be obtained in this case.

5. Discussion.

The experiments show that the behaviour of swollen cellulose gels upon deformation and subsequent drying in its general features conforms to what should be expected from a molecular network structure whose degree of coiling in the primary gel is independent of the volume of the primary gel and whose orientation remains practically constant upon drying in the anisotropic state. This implies that orientation as a function of v_t is more rapid according as the degree of swelling of the primary gel is lower. The fact that the slope of the anisotropy against v_t curves times $q_0^{2/3}$ give somewhat higher values in cellulose gels than in xanthate gels is not surprising since the constant k in equations (2), (4) and (5) includes factors which depend on the stiffness of the molecular chains (length and anisotropy of Kuhn's "statistical chain element").

There are, however, other arguments showing that the picture cannot be more than a first and rough approximation. To begin with the experimental anisotropy curves (Fig. 2) remain straight to much higher values of v_t than can be expected from equation (5) and further it can be shown that the absolute magnitude of their slopes cannot be accounted for by Kuhn's theory which assumes a large number of chain-elements between the junction points of the network. This would lead to much smaller slopes. It can be shown that the slope of the curves must increase according as the number of chain elements per chain decreases and an estimation of the absolute values leads to the very small number of 1 or 2 chain-elements in these cellulose gels. For chains of this length Kuhn's theory is, of course, no longer applicable. Furthermore, the rapid decrease in volume upon unidirectional stretching is not in conformity with long chains, which should rather give rise to a volume increase. This points to a network of a much more rigid character.

Various attempts have recently been undertaken to develop the theory also for molecular networks with shorter [16] and very short chains amongst which we particularly wish to mention that by J. J. Hermans,[17] which seems to be capable of a quantitative interpretation of certain important features of the experimental behaviour (see the following paper), though leading to some difficulties of another kind.

Finally it should be remarked that the influence of the crystalline part of the objects studied, which roughly amounts to 20 %,[18] has not been mentioned nor accounted for in this paper. The rate of orientation of the crystallites as investigated by X-ray analysis will be dealt with elsewhere.[19]

Cellulose Res., A.K.U., and Affiliated Co.,
Utrecht,
Holland.

---

[16] James and Guth, *J. Chem. Physics*, 1943, **11**, 455; *J. App. Physics*, 1944, **15**, 294. Treloar, *Trans. Faraday Soc.*, 1946, **42**, 83.

[17] J. J. Hermans, *J. Colloid Sci.*, 1946, **1**, 235.

[18] P. H. Hermans, *Contribution to the Physics of Cellulose Fibres* (Elsevier, Amsterdam-New York, 1946).

[19] P. H. and J. J. Hermans, Vermaas and Weidinger, *J. Polymer Sci.* (in press)

THE STRUCTURE AND THE MECHANISM OF DEFORMATION IN CELLULOSE GELS.

II. APPLICATION OF THE THEORY TO SHORT CHAINS.*

By J. J. Hermans.

Received 13*th August*, 1946.

A theory for networks consisting of short chain fragments given in a previous paper[1] applies Kramers' method,[2] which was developed for chains of arbitrary length. In two respects it differs from that discussed in Part I. In the derivation of equations (1) to (6) of Part I it is assumed that the junction point pattern follows the macroscopic changes in the dimensions of the bulk substance, and that the number N of statistical chain elements between two successive junction points is large. The theory referred to is not restricted to long chain fragments and, instead of dictating certain displacements to the junction points, it calculates the effect of an overall force on the chains in the direction of the stretch.

1. Force-strain Relation.

The results of this theory are most conveniently given in terms of a parameter λ which is proportional to the stretching force. If K is the force per unit area of the original cross-section, G_0 the number of chains in unit volume of original unstretched gel, $\bar{z}_0$ the average projection of the chains on the direction of stretch in the original unstretched gel, and A the length of a statisical chain element in Kuhn's sense, we have

$$K = \lambda kT\, G_0 \bar{z}_0 / A \quad . \quad . \quad (1)$$

The theory gives the degree of stretch v in terms of λ for different values of N. This relation between v and λ is represented in Fig. 1 for $N = 1, 2, 3$. The experimental force-strain curves, however, are different from these theoretical curves in that they show a yield value OY (compare the dotted curve in Fig. 1). This yield value is very small for highly swollen gels, but reaches quite perceptible values at lower degrees of swelling. When applying the theoretical force-strain relations of the type shown in Fig. 1, we must first reduce the experimental curves to zero yield value. We have done this by simply subtracting the yield value observed from all the force values on the force-strain curve. In other words, the dotted curve in Fig. 1 is simply shifted towards the origin

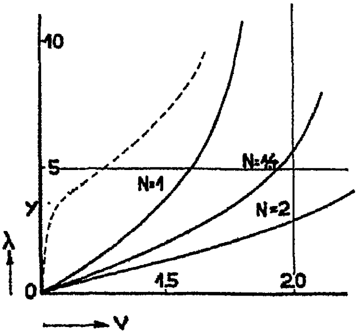


FIG. 1.—Relation between degree of stretch v and force parameter λ. Dotted curve : experimental (λ in arbitrary units).

---

* Communication No. 41 from the Laboratory for Cellulose Research, AKU and Affiliated Companies, Utrecht, Holland.

[1] J. J. Hermans, *J. Colloid Sci.*, 1946, 1, 235.

[2] Kramers, *Physica*, 1944, 11, 1 ; *J. Chem. Physics* (in press).

over a distance OY in the direction of the force axis. This method is admittedly of a somewhat arbitrary nature. Its physical meaning is, that we attribute the yield value to a mechanism which has no direct bearing on the general form of the curve.

Doing so, the experimental force-strain curves are found to be strikingly similar to the theoretical curves. In fact, if we take N between 1 and 2, the ratio between the experimental force K and the theoretical parameter λ is constant within experimental error in the entire elongation range from 0 to 100 % stretch. According to equation (1), the value of this constant ratio determines the magnitude of $G_0\bar{z}_0/A$. Here $\bar{z}_0/A$ has a value somewhat smaller than 1. To be exact, if coiling is "normal", $\bar{z}_0/A = 0{\cdot}5$ for $N = 1$ and $0{\cdot}7$ for $N = 2$ (compare ref. 1), and coiling cannot alter the order of magnitude of $\bar{z}_0/A$. Consequently, we are in a position to estimate G_0, the number of chain fragments in unit volume. Since we know the density of the polymer substance and the molecular weight of the monomer, we can calculate the number of monomer groups in the chain fragment between successive junction points, and thus also the number of monomer groups in the statistical chain element; the result is about 6 monomer groups per chain element. This is roughly three times as small as the value derived from viscosity data.[3] It may be mentioned in this connection, however, that the force-strain curves depend on the rate of stretching. From relaxation experiments at constant elongation it follows that all forces are reduced to about half their value if the rate of stretch becomes very small.[4] If this reduced force-strain curve is compared with theory, the agreement is equally good, but the value of G_0 becomes about twice as small and the number of monomer units per chain element becomes about twice as large.

2. Relation between Force and Anisotropy.

If we apply equation (1) to a series of gels of the same origin (i.e. equal degrees of swelling, q_0, in the primary gel) but with different degrees of swelling q_1, the value of $\bar{z}_0$ in these gels will be proportional to $q_1^{1/3}$. Since G_0 is, of course, proportional to q_1^{-1}, the force K for a given value of λ will be proportional to $q_1^{-2/3}$. On the other hand, it can be shown that for N-values below about 4, the relation between force parameter λ and anisotropy of the fibre is practically independent of N and is given by a slightly sigmoid-shaped curve. By anisotropy in this context we mean the so-called orientation factor f, which measures the ratio between the double refraction observed and that at complete orientation. Consequently, "for a given value of λ" amounts to the same thing as "for a given value of the orientation factor". In other words, if we compare two gels of the same origin, a plot of $Kq_1^{2/3}$ against anisotropy f for the one gel must coincide with the same plot for the other gel. This is borne out by Fig. 2, which also confirms the sigmoid shape.[5] The three gels compared are: (1) xanthate fibre,

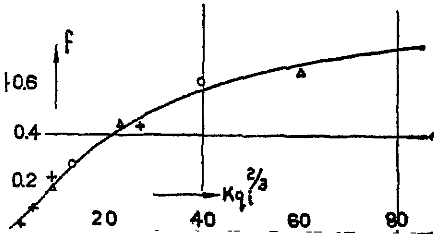


FIG. 2.—$Kq_1^{2/3}$ plotted against orientation factor f; (o) xanthate gel, (Δ) fresh cellulose gel, (+) reswollen cellulose gel. (Viscose No. 5 of the preceding paper).

---

[3] J. J. Hermans, *Kolloid Z.*, 1944, **106**, 22; *Rec. trav. chim.*, 1944, **63**, 25.

[4] J. J. Hermans, *ibid.* (in press)

[5] Strictly speaking, we should insert the anisotropy of the amorphous part of the fibre, which is to be distinguished from that of the crystalline part. The latter can be derived from the X-ray diagram (P. H. and J. J. Hermans, Vermaas

$q_1 = 13$, (2) the same fibre decomposed to cellulose, $q_1 = 5{\cdot}9$ and (3) this cellulose gel after drying in air and reswelling in water, $q_1 = 2{\cdot}3$.

3. Volume Changes.

As was pointed out in Part I, the decrease in volume observed when stretching highly swollen xanthate gels or fresh cellulose gels cannot be explained on the basis of networks containing long chain fragments (i.e. large number N of statistical chain elements between successive junction points), because for large N-values the theory requires an increase in volume on stretching.[6] It is obvious, however, that even in this case the volume increase is restricted to moderate elongations. At more advanced degrees of stretch any network will eventually close its meshes and squeeze out the swelling medium. It depends on the value of N, where this effect becomes predominant; on stretching a network of stiff rods ($N = 1$), the volume is decreased from the very beginning.

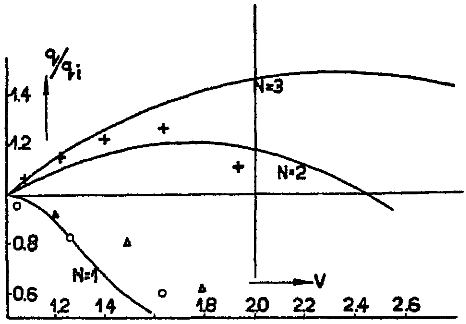


FIG. 3.—Relative volume q/q_1 in dependence on stretch, according to theory (drawn curves) and experiment. (o) cellulose xanthate, (Δ) fresh cellulose, (+) reswollen cellulose (Viscose nr. 5).

A rough estimate of the volume change for low values of N can be derived from theory. To that end we calculate the average value of $|\sin\theta|$, where θ represents the angle between the vector joining the end points of a chain and the direction of the stretch. The calculation can be carried out according to the method developed in reference (1). It is obvious that $|\sin\theta|$ gives a measure of the linear dimensions of the gel in directions perpendicular to the stretch. It is to be noted that this calculation ignores the influence which the free energy of dilution will doubtless have on the volume change [6, 7]; it treats the problem from the purely mechanical viewpoint: the gel frame is subjected to a stretching force and responds to this force by changing its configuration and volume; the "thermodynamical" response of the two-component system to this volume change is not taken into account.

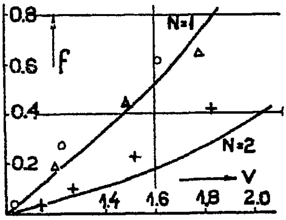


FIG. 4.—Anisotropy (= orientation factor) f of xanthate (o), fresh cellulose (Δ) and reswollen cellulose (+) compared with theoretical curves for $N = 1$ and 2 (Viscose nr. 5).[8]

Fig. 3 gives the theoretical curves for the volume change as a function of elongation v. Comparing with the experimental points for cellulose xanthate, fresh cellulose and reswollen

---

and Weidinger, *Rec. trav. chim.* (in press), the former is calculated from the total anisotropy and the crystal anisotropy and requires the knowledge of the amount of crystalline material in the fibre. For the present, however, we shall restrict ourselves to considerations of a more general nature and simply insert the overall anisotropy of the stretched gel.

[6] Flory, *J. Chem. Physics*, 1943; J. J. Hermans, *Trans. Faraday Soc.* (in press).

[7] Flory and Rehner, *J. Chem. Physics*, 1943, **11**, 521.

[8] Here again, we have simply inserted the overall anisotropy of the fibre. The experimental points compare somewhat better with the theoretical curves if we insert the anisotropy of the amorphous part.

cellulose, which are also shown in the figure, we observe that this simple theory is capable of giving a fairly correct description of the phenomena studied if we assume that $N = 1$ in xanthate and fresh cellulose gels, whereas $N = 2$ in reswollen cellulose. This is in conformity with the relation between anisotropy and elongation as illustrated by Fig. 4.

4. Conclusion.

The behaviour of cellulose gels can be explained to a first approximation by assuming a network of short chain fragments. The number N of statistical chain elements between successive junction points is hardly more than 1 in cellulose xanthate and fresh cellulose and about 2 in reswollen cellulose. Presumably this increase in N as a result of the lower degree of swelling may be considered as the equivalent of " super-coiling " in a network where N is large, although it is more difficult to picture such a " super-coiling " in a network of stiff rods. We must expect, therefore, that the higher the degree of swelling of the primary gel compared with that of the reswollen gel, the more pronounced this "super-coiling ". This is expressed by a somewhat higher value of N in this reswollen state, and accordingly by a higher maximum in q/q_1 at a somewhat higher v-value. This is completely borne out by experiment.

It is obvious that the theory can do more no than give a very rough first approximation. For example, the value $N = 2$ in reswollen cellulose represents only an average value, the finite thickness of the chain elements is neglected, the stretching force is not distributed equally over all chains, etc. All these approximations will become more and more important as the degree of stretch is increased. However, the general properties of cellulose gels on deformation are fairly satisfactorily described by the present theory.

The experiments were carried out by Dr. D. Vermaas of this Institute.

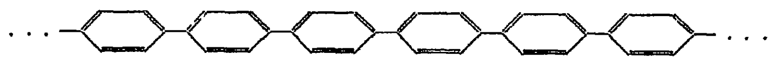
Cellulose Res., A.K.U., and Affiliated Co.,
Utrecht,
Holland.

GENERAL DISCUSSION *

Dr. R. J. B. Marsden (*Manchester*) said : Much has been said of the " coiling " and " uncoiling " (*pace* Treloar) of long molecules during solvation and swelling. Surely on *a priori* or at least ordinary stereochemical grounds, the rigidities of some of the molecular types discussed are of entirely different orders ? Cannot this rigidity be arrived at, even approximately, by making use of the known radii of atoms when both combined and when merely pushed into contact ? I believe Sidgwick [1] suggested (he has certainly used) an arbitrary additional radius of about 0·5 A. for this purpose, and there are the Stuart models.

Thus the ability to coil of such molecules as (*a*) polythene and rubber, (*b*) the jammed structure of poly*iso*butylene, (*c*) starch (with its ability to form hollow helices of six anhydro-glucose units) and (*d*) cellulose, which in a model appears to be very rigid, must have rigidities and coil sizes which are very different and for which at least approximate values could be obtained from known atomic dimensions and molecular structures.

If the organic chemists could provide us with, say, a poly-diphenyl, thus :



it should be exceedingly rigid and would probably be a useful experimental material, particularly for viscosity studies.

---

* On five preceding papers.

[1] *Ann. Rep.*, 1932, 70.

There certainly now seems room for the synthesis of special molecules with pre-determined properties of this sort purely as research materials, instead of those which are only of technological value.

Mr. Warburton (*Leeds*) (*communicated*): In addition to the methods of obtaining the transverse swelling of fibres mentioned by Dr. Preston, reliable values can be obtained from density measurements in benzene, etc., when the longitudinal swelling is known. The writer has used the density of King,[2] to obtain the transverse swellings of wool fibres and compared the results with other methods.[3] It has also been used by Speakman and Saville to obtain the transverse swelling of Nylon.[4] There are many references to density measurements in the literature and most of these are given by Hermans.[5]

Dr. A. R. Urquhart (*Lancaster*) said: Referring to sorption hysteresis, Dr. Barkas states that "from the fact that no hysteresis occurs in true solutions, where no shear strains are possible, we may infer that plasticity is the cause of sorption hysteresis in gels." Recently Speakman and Saville [6] showed that there is no hysteresis in the sorption of water vapour by Nylon, and, in terms of an earlier explanation of hysteresis put forward by myself, attributed its absence to the fact that in Nylon the main chains carry no lateral hydrophilic groups. It would be of interest to learn how Speakman's observations fit into Barkas' scheme.

Mr. F. A. P. Maggs (*London*) (*communicated*): The connection between swelling and hysteresis of the adsorption isotherm to which Dr. Barkas has drawn attention is a fairly widespread phenomenon for solids which show a *large* (say, greater than 2 %) adsorption swelling; the deformation for such solids is most probably of a non-elastic nature. In the case of solids which swell but little, however, elastic deformations are more frequently encountered and other explanations of hysteresis must be sought; there is, for instance, evidence to show that hysteresis is often observed when phase changes in the adsorbed film take place.[1] Considerable caution must therefore be exercised in attributing hysteresis in general to a cause which probably operates only in a limited number of systems.

Dr. J. St. L. Philpot (*Maidenhead*) said: In answer to Dr. Marsden's suggestion that rigid polymers should be synthesised, I should like to mention that some work on these lines is being done at Maidenhead.

Dr. Barkas (*Princes Risborough*) (*communicated*): Dr. Urquhart will find my views on the absence of hysteresis in Nylon and its presence in natural fibres briefly expressed in my comments on the paper by Hailwood and Horrobin. To explain hysteresis, my theory requires the existence of shear stresses in a hydroscopic material possessing plasticity of a type where the stress does not completely disappear under plastic flow. Accordingly, no hysteresis is predicted for the sorption either of true solutions, however viscous, or of perfect crystals, but it would be expected in any system where the free sorption on large molecules is impeded by irregular cross-linkages between fairly rigid groups of molecules, or by restraints introduced by the anatomical structure of fibres. If the main chains of Nylon carry no lateral hydrophilic groupings, there must be little tendency for shear stresses to result in the gel itself and if, in drawing Nylon fibres under tension, some molecular orientation results, these will lie mainly parallel to the fibre axis, an arrangement which will not set up appreciable shear stresses on subsequent swelling and will certainly be less effective than the restraints in natural wood fibres, for example, where the swelling of a hollow cell is reduced to about one half of what one would expect under free sorption of the gel material.

[2] King, *J. Text. Inst.*, 1926, **17**, T53. [3] Warburton, *ibid.*, 1946 (in press).

[4] Speakman and Saville, *ibid.*, 1946, **37**, P279.

[5] Hermans, *Contribution to the Physics of Cellulose Fibres* (Amsterdam, 1946), p. 73.

[6] Speakman and Saville, *J. Text. Inst.*, 1946, **37**, P271.

Mr. Maggs rightly sounds a note of warning against a too wide application of my theory of hysteresis, and, if I have appeared too optimistic in this paper, I welcome his correction. If, however, he will refer to my first detailed account of the theory [7] he will find that I have been suitably cautious and on page 44, ref. 1 he will see that I have accepted Gregg's hypothesis of phase changes in adsorbed films.

[7] *Trans Faraday Soc.*, 1942, 38, 15.

ON THE STRUCTURE OF TRANSPARENT WATER AND OIL DISPERSIONS (SOLUBILISED OILS).

By J. H. Schulman and T. S. McRoberts.

Received 7th August, 1946.

Hoar and Schulman [1] have presented some semi-quantitative evidence that transparent systems of oil and water were made up of so-called swollen water micelles of dimensions of 120-200 A. diameter, and that these water droplets were dispersed in the non-conducting oil phase. Since the water droplets were below the dimensions of 1000 A. ($\frac{1}{4}\lambda$ of light) the dispersions appeared transparent. The necessary characteristics of these transparent oil and water dispersions containing equal quantities of oil and water were high soap concentrations and the presence of an unionised amphipathic molecule such as an aliphatic alcohol or cresol, or amine, etc., in at least equimolecular concentrations with the soap. Addition of water to the non-conducting systems seems to swell these droplets to visible dimensions (Tyndall effect), still keeping the system non-conducting. Further addition of water inverted the system to an oil-in-water opaque emulsion. It was suggested that the droplets were stabilised by an interfacial mixed monolayer of soap and alcohol. Recent work by Stauff, Hess, Kiessig, Philippoff, McBain and Harkins,[2] has shown, by the X-ray study of soap solutions, that the soap micelle has an ordered lamellar molecular structure which can be swollen within certain limits both by water and oil.

Since these transparent oil and water systems contain oil in proportions many times greater than the quantity of soap, and since the dimensions of the droplets would appear to approach those of the swollen soap micelle, it was considered important to study the role the alcohol molecules play in enabling the aqueous soap system to absorb large quantities of oil and how they controlled oil or water phase continuity.

Experimental.

In a system always containing equal amounts of oil (benzene, paraffin, Nujol, etc.) and water, soap was added to the water phase in specified concentrations. Sodium oleate was chosen as the soap to use in these investigations since at 20° C. it is definitely above the Krafft point. It has the disadvantage of containing a

[1] Hoar and Schulman, *Nature*, 1943, 152, 102.

[2] (*a*) Hess and Gundermann, *Ber.*, 1937, 70, 1800. (*b*) Philippoff and Hess, *ibid.*, 1937, 70, 1808. (*c*) Hess, Kiessig and Philippoff, *Naturwiss.*, 1938, 26, 184. (*d*) Hess, Philippoff and Kiessig, *Kolloid Z.*, 1939, 88, 40. (*e*) Stauff, *ibid.*, 1939, 89, 224; *Naturwiss.*, 1939, 27, 213. (*f*) Kiessig and Philippoff, *ibid.*, 1939, 27, 593. (*g*) Kiessig, *Kolloid Z.*, 1941, 96, 252. (*h*) Philippoff, *ibid.*, 1941, 96, 225. (*i*) Kiessig, *ibid.*, 1942, 98, 213. (*j*) Hess, *Fette Seifen*, 1942, 49, 81. (*k*) Hughes, Sawyer and Vinograd, *J. Chem. Physics*, 1945, 13, 131. (*l*) Harkins, Mattoon and Corrin, *J. Amer. Chem. Soc.*, 1946, 68, 220. (*m*) Ross and McBain, *ibid.*, 1946, 68, 296.

small percentage of straight chain soaps as impurities, but these give transparent oil-water systems similar to those given by sodium oleate. These opaque viscous masses were titrated to fluid transparency with normal alcohols varying in chain length from C_1 to C_{10}, and also with cyclic alcohols such as cyclo-hexanol, methyl cyclohexanol, phenol and cresol. The end-point of fluid transparency is easily recognised, and is sharp with change in concentration of alcohol. The amounts of oil and water were equal and constant for the whole experimental system. The conductivity as measured with two spaced copper wires attached to an avometer potentiometer box denoted very readily whether oil or water was the continuous phase. If oil was the continuous phase, no conductivity was obtained. Owing to the soap concentrations used in the water phase, it was easy to determine by their high conductivity whether the aqueous phase was the continuous one.

TABLE I.

Amphipathic Compound	Phase Continuity		
Chain Compounds	Benzene	Paraffin (B.P. 180–220° c.)	Nujol (B.P. 160–200° c. at 1mm. Hg.)
Ethyl alcohol	Water	No dispersion	No dispersion
Propyl alcohol	Water	No dispersion	No dispersion
Butyl alcohol	Water	Water	No dispersion
Amyl alcohol	Water	Water	Water
Hexyl alcohol	Oil	Oil	Water
2-Ethyl hexanol	Oil	Oil	Oil
Decyl Alcohol	Oil	Oil	Oil
Cyclic Compounds			
C_6 Cyclohexanol	Water	Water	Water
C_6 Phenol	Water	Water (unstable)	Water (unstable)
C_7 *p*-Methyl cyclohexanol	Oil	Water	Water
C_7 *m*-cresol	Oil	Water	Water

It was found very useful to use copper wires as the electrodes, since the copper surface is hydrophilic. Trouble is always experienced with platinum electrodes which can become contaminated by oil.

The systems usually contained 5 cc. water, specified sodium oleate concentration and 5 cc. benzene, to which the alcohol is titrated. Experiments were carried out at room temperature (approximately 20° c.).

Results.

(a) Oil or Water Phase Continuity in Relation to the Number of Carbons in the Alcohol.—Ethyl, propyl, butyl and amyl alcohols give conducting systems with benzene, over the whole range of the soap concentration. The continuous phase inverts to the non-conducting oil very sharply at hexyl alcohol, and for all the higher alcohols, i.e. inversion takes place between C_5 and C_6 carbon atoms in the amphipathic molecule. Cyclohexanol gives a water-continuous system, but *p*-methyl cyclohexanol an oil-continuous system; phenol also gives a water-continuous, but *m*-cresol an oil-continuous system.

These systems exist when benzene is the oil phase, but when a more hydrophobic oil such as medicinal paraffin (thin Nujol) is used, the change over takes place at C_7 to C_8 carbon atoms in the amphipathic dispersing molecule. The water phase is the continuous one for all the cyclic alcohols in the presence of the paraffins (see Table). An explanation of these results will be brought out in the latter part of this paper.

It was found that varying temperature had no effect on phase inversion.

(b) Oil and Water Dispersion in Relation to Alcohol and Soap Concentration.—In Fig. 1 is given the quantity of an alcohol necessary to give a fluid transparent system for different concentrations of soap, the oil-water ratio remaining constant and equal. Propyl alcohol gives a transparent droplet system to comparatively low soap concentrations with a breakdown into two phases at low soap concentration. In the range just before breakdown, strong Tyndall effects were observed. At very low soap and high propyl alcohol concentrations there is a solution effect on a quaternary system as with ethyl alcohol.

n-Butyl, amyl, hexyl and decyl alcohol systems appear to give true transparent droplet systems with sodium oleate, but with the interesting characteristic that with increasing chain length of the normal alcohol, the system breaks down at higher minimum concentrations of soap. Butyl alcohol gives an anomaly at low soap concentrations ; this might be due to the fact that the product of mixed solubility of butyl alcohol in benzene and water is lowest for all the alcohols

In the cyclic alcohol series a very interesting anomaly is observed with methyl cyclohexanol. The addition of the methyl group to cyclohexanol enables a transparent system to be obtained at low concentrations of soap. At these low soap concentrations strong Tyndall effects appear, until at the lowest concentration a very fine emulsion is observed ; as will be discussed later in this paper, there appear to be differences in the interfacial order and disorder according to whether the penetrating molecule is cyclic or straight chain.

It is interesting to note that with a long chain alcohol such as decyl alcohol the breakdown in the system occurs at high soap concentration and at an equi-molecular mixture of the soap-alcohol.

(c) The effect of variation of the Oil/Water Ratio.—It was shown that the system as shown *in* Fig. 1 remains stable for wide variations in the oil/water ratio at high soap concentrations, and no phase change takes place. On increasing the oil/water ratio in a water continuous system to large values, keeping the total volume of oil and water constant, a gradual change in conductivity is observed, over the range 90 to 95 % oil. This system is non-conducting at 95 % oil. On decreasing the oil/water ratio in an oil continuous system to small values, keeping the total volume of oil and water constant, a gradual increase of conductivity is observed, starting at 20 % oil.

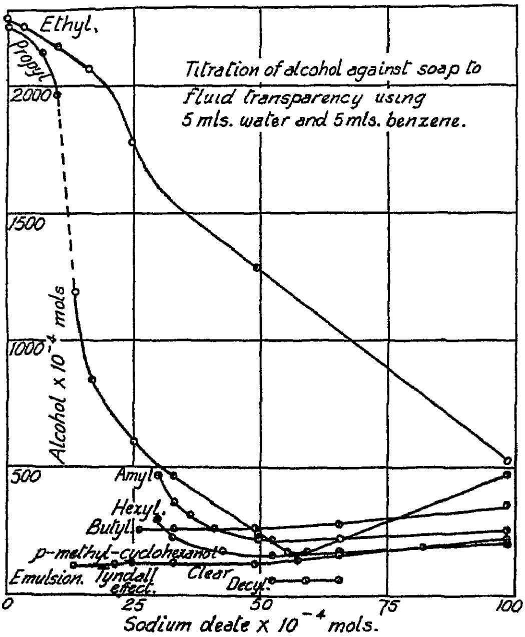


FIG. 1.

***(d)* X-Ray Analysis.**—*The* X-ray structures *of the* systems so far prepared are being investigated.

Discussion.

(a) Phase Inversion.—It can be seen from the phase inversion experiments (Table) that the number of carbons, and to some extent their arrangement, either cyclic or straight chain in the amphipathic molecules, is responsible for organising the overall continuity of the oil or water phase. This suggests that the amphipathic molecules must sit with the soap molecules at an interface and change the wettability of the interface by the oil and water according to the balance of the polar portion and non-polar portion of the associating amphipathic molecule.

Thus, with a straight chain of 5 carbons or below, or a cyclic hydrocarbon of 6 carbon atoms in the associating molecule, saturated or unsaturated, the contact angle of the interfacial layer must be smaller with water than benzene, thus making water the continuous phase. On increasing the chain length or addition of methyl groups to cyclic hydrocarbons, the contact angle must be smaller with benzene, thus making benzene the continuous phase—see Fig. 2. (These are the same rules for phase inversion as described for stabilisation with solid particles).[3]

It can be seen from the Table in general that the hydrophobic-hydrophilic balance in the structure of the interface is important for controlling the continuity of the oil or water phase for the same oil. The relative wettability of the different oils by the interface can likewise govern the phase continuity and be responsible for the possibility of the formation of a mixed interfacial film. Thus it can be seen that benzene, but not paraffin or Nujol can wet a mixed propyl alcohol-sodium oleate interface sufficiently to form a dispersion. The addition of one carbon to the interfacial film enables paraffin to wet it so as to permit a paraffin-in-water dispersion with butyl alcohol. Further addition of one carbon, as with amyl alcohol enables Nujol to be dispersed in water. This suggests that not only has the soap lamellar film to be penetrated by an amphipathic molecule to enable droplets to be formed, but the structure of the mixed film itself must be compatible with the oil phase, or the formation of a mixed interfacial film will not occur.

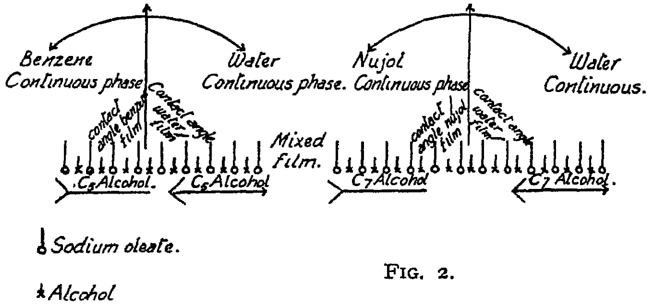


FIG. 2.

(b) Order and Disorder in the Interfacial layer.—It could be further suggested from this evidence that the soap molecules in the orientated layers of the soap micelle cannot swell beyond a certain extent, unless the crystal arrangement of the soap molecules in the lattice is somewhat broken down by the penetration of a molecule which could create a certain disorder in the lattice. This disorder would cause the solid structure to break down and permit a liquid or highly viscous interfacial film to form. The liquid interfacial film could thus expand and contract and permit the formation of droplets containing oil or water according to the wettability of the interfacial film. Assuming that all the soap is at the interface and that the area per molecule of this partially or fully penetrated soap film is 40 A^2. per sodium oleate molecule, the diameter of these droplets can be calculated for each soap concentration. It is interesting that in the case of *p*-methyl cyclohexanol, which permits water/oil dispersions down to comparatively low concentrations of soap, a strong Tyndall effect starts in at the calculated diameter of 800 A. A fine brown emulsion denoting a diameter of 1000 A. for the droplets begins at the expected soap concentration (see Fig. 1).

Further support for this theory of disorder in the micelle crystal lattice, permitting a liquid interface to form and hence permitting droplets to occur, is given by the fact that penetration of the interface is noticeable at C_3 carbons in the penetrating molecule. This chain length is where one would first expect penetration of the interface to be noticeable. On increasing the chain length, a greater order is being brought back to the interfacial layer and it becomes more difficult to form dispersions of oil and water. This becomes very marked with decyl alcohol where equimolecular proportions of alcohol and soap are observed for the minimum concentration of soap which allows a dispersion. On increasing the chain length further one would expect to obtain a mixed solid crystal lattice again and no swelling other than micelle swelling would take place. This is indicated by the fact that with increasing chain length of alcohol a higher minimum concentration of soap is necessary to bring about oil and water dispersions.

---

[3] N. K. Adam, *The Physics and Chemistry of Surfaces*, 2nd ed. (Oxford), p. 207.

It is interesting that with the methyl cyclohexanol, the influence of the ring structure is to produce maximum disorder of the sodium oleate lattice compatible with stability of the mixed interfacial film. The sudden increase in the quantity of alcohol used at high soap concentrations, especially for the small alcohol molecules, is necessary to peptise the soap which is surplus to that required to form the interfacial film.[4]

One can thus visualise the structure of these soap-oil-water dispersions as commencing with a lamellar micelle in strong soap solution. This micelle can be partially swollen with oil without breakdown of the crystal lattice structure. This lattice can be penetrated by small amphipathic molecules, permitting the interface to become liquid and thus permitting droplets to form. On decreasing the soap concentration, the droplets must swell to ensure a decreasing interfacial area. Where the penetrating molecule has strong adhesion to the soap molecule, the interfacial film breaks down at an equimolecular mixture of the alcohol-soap. In this case, owing to the fact that a certain order has been brought back to the interfacial film, the droplets cannot be swollen in low soap concentration (decyl alcohol cannot stabilise a droplet below 32 % soap on the water). But with the smaller penetrating molecules, and especially methyl cyclohexanol, the droplets apparently swell on decreasing the soap concentration until they become visible, first as in a Tyndall beam and then as very fine emulsions. The diameters of the droplets are quite close to those calculated, assuming most of the soap is adsorbed at the oil/water interface with an area per molecule of *ca.* 40 A^2. per soap molecule. This value of 40 A^2. per sodium oleate molecule is deduced from the fact that the side spacing in the X-ray pictures gives 4·5 A., which would give a packing in the micelle of 20 A^2. per molecule. A fully penetrated sodium oleate layer by a straight chain molecule would give an area of 40 A^2. per sodium oleate molecule. This consideration is somewhat different from the dimensions obtained by considering a sodium oleate molecule at an oil/water interface in dilute solution, where the molecule is dissociated instead of being in the form of an ion pair as in strong soap solutions. The dissociated molecules repel one another and the area per sodium oleate molecule is given as 30 A^2. Insertion of an hydroxyl dipole between each carboxyl would again give an area per molecule of sodium oleate of approximately 40 A^2.

An interesting feature of the penetration of the soap lattice by alcohol molecules is that when the hydrocarbon chain is straight, as in hexyl or decyl, much smaller quantities of the alcohol are necessary to bring about dispersion than if the chain is branched. This is usual in monolayer penetration experiments, where, if steric hindrance is experienced in adlineation between two molecules, the surface pressure in the mixed film is reduced.[1] Much higher concentrations of the branched chain penetrating molecule are required to establish identical surface pressures as compared with a solution of straight chain molecules.

Hexyl alcohol required 2½ times less concentration than 2-ethylhexanol to enable sodium oleate to disperse water in paraffin, whereas much less decyl alcohol is required than either of the two alcohols.

---

[4] Lawrence, *Trans. Faraday Soc.*, 1937, **33**, 325; see also Pink, *J. Chem. Soc.*, 1939, 53.

Summary.

To equal volumes of oil and water, soap in varying concentrations has been added. This mixture has been titrated to fluid transparency by phenols and aliphatic alcohols of varying chain length. It was found that the number and arrangement of the carbon atoms, both in the alcohol and in the oil, decided whether the system was conducting or non-conducting. The explanation given for this phenomenon is that the continuity of the phase is related to the contact angle between the oil, water and the mixed interfacial film of soap/alcohol. The suggestion is given that the alcohol molecule penetrates the lamellar lattice of the swollen soap micelle permitting droplets of oil or water to form by liquefaction of the interfacial film. These droplets can swell, on decreasing the soap concentration, to visible dimensions according to the associating and dispersing powers of the alcohol molecule with the sodium oleate molecule.

We wish to thank Prof. E. K. Rideal, F.R.S., for stimulating help, and Queen's University, Belfast, for the award of a Musgrave Studentship (T. S. McR.).

The Department of Colloid Science,
The University,
Cambridge

CRITICAL EFFECT OF TEMPERATURE ON THE ABSORPTION OF WATER BY SOLUTIONS OF ETHANOLAMINE OLEATE IN BENZENE.

By R. C. Pink.

Received 14th October, 1946.

A recent paper by Schulman and McRoberts [1] has provided a very suggestive interpretation of the well-known action of phenols and alcohols in increasing the amount of water dissolved by solutions of soaps in organic solvents. According to this interpretation based on the work of Stauff, Hess, Kiessig, Philippoff, McBain and Harkins,[2] the soap micelle cannot swell beyond a certain extent unless the crystalline arrangement of the molecules in the lattice is broken down by the penetration of an un-ionised amphipathic molecule which will create sufficient disorder in the lattice to permit a liquid interfacial film, composed of soap and the amphipathic molecule, to form. This liquid film, which may be highly viscous, can expand and permit the formation of droplets of sizes up to visible dimensions. In the experiments of Schulman and McRoberts [1] and in earlier work,[3] a necessary characteristic of the systems containing equal quantities of water and oil appeared to be the presence of an aliphatic alcohol or phenol in at least equimolecular concentration with the soap.[4]

---

[1] This discussion.

[2] (*a*) Hess and Gundermann, *Ber.*, 1937, **70**, 1800; (*b*) Philippoff and Hess, *ibid.*, 1937, **70**, 1808; (*c*) Hess, Kiessig and Philippoff, *Naturwiss.*, 1938, **26**, 184; (*d*) Hess, Philippoff and Kiessig, *Kolloid Z.*, 1939, **88**, 40; (*e*) Stauff, *ibid.*, 1939, **89**, 224; *Naturwiss.*, 1939, **27**, 213; (*f*) Kiessig and Philippoff, *ibid.*, 1939 **27**, 593; (*g*) Kiessig, *Kolloid. Z.*, 1941, **96**, 252; (*h*) Philippoff, *ibid.*, 1941, **96**, 225; (*i*) Kiessig, *ibid.*, 1942, **98**, 213; (*j*) Hess, *Fette Seifen*, 1942, **49**, 81; (*k*) Hughes, Sawyer and Vinograd, *J. Chem. Physics*, 1945, **13**, 131; (*l*) Harkins, Mattoon and Corrin, *J. Amer. Chem. Soc.*, 1946, **68**, 220; (*m*) Ross and McBain, *ibid.*, 1946, **68**, 296.

[3] Hoar and Schulman, *Nature*, 1943, **152**, 102.

[4] See also Weicherz, *Kolloid Z.*, 1929, **47**, 133; 1929, **49**, 158.

It is clear, however, that if the interpretation above is correct, the necessary disorder in the micelle could, in principle, be produced by increase in temperature alone. Provided, therefore, a suitable soap is chosen, similar transparent systems containing large quantities of water relative to the volume of oil should be obtainable in the absence of alcohols or phenols, merely by raising the temperature. Experimental evidence for this has been obtained in the investigation, described below, into the absorption of water by ethanolamine oleate dispersions in benzene. This was carried out in 1939 but has not hitherto been published.

Experimental.

Known amounts of water were added to 50 cc. of the benzene/ethanolamine oleate solution and the temperature varied until the point was determined where a slight decrease in temperature caused the solution to become faintly cloudy. In order to detect this point with sufficient accuracy the solution was contained in a Pyrex tube supported in a large water-bath through which a strong beam of light was passed. When the temperature of the bath was varied very slowly the point at which the solution became cloudy could be detected with an accuracy of 0·5° c. The results obtained with soap solutions of different concentrations are shown in Fig. 1. Below a critical temperature, in this case 50° c., the addition of excess water produces a curdy precipitate of hydrated solid soap. In the neighbourhood of 50° c. the amount of water absorbed by the solutions increases in a remarkable manner and, instead of a precipitation of soap, the addition of water gives rise in the first place to a faint Tyndall effect. With the stronger soap solutions this is followed firstly by a greenish opalescence and eventually by the formation of a finely divided emulsion. In the dilute soap solutions an emulsion is formed when only small quantities of water have been added. The transparent systems produced above 50° c. are exactly similar in appearance to those produced at room temperature when large amounts of water are added to solutions of ethanolamine oleate in benzene to which phenols or cresols have been added.[5]

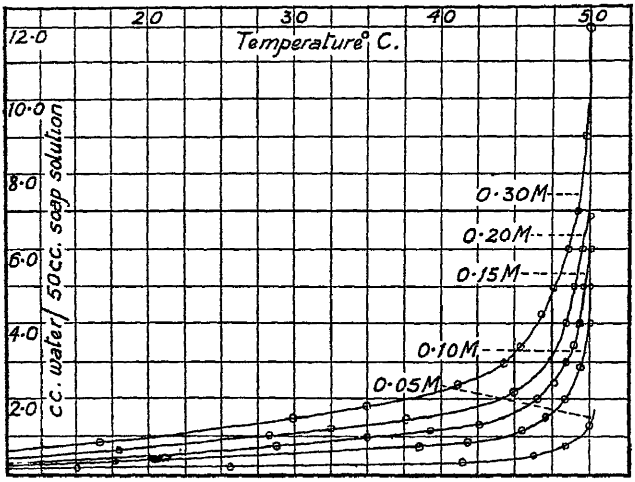


FIG. 1.—The critical effect of temperature on the absorption of water by benzene solutions of ethanolamine oleate.

---

[5] Pink, *J. Chem. Soc.*, 1939, 53.

Viscosity Measurements.—In order to determine whether any major change in the degree of aggregation of the soap occurred in the critical temperature region around 50° c., the viscosities of the soap solutions were measured by the Ostwald U-tube method. Measurements were made with the dry soap solutions and also with solutions to which a certain amount of water had been added. The results obtained with 0·1 and 0·3 M soap solutions are shown in Fig. 2.

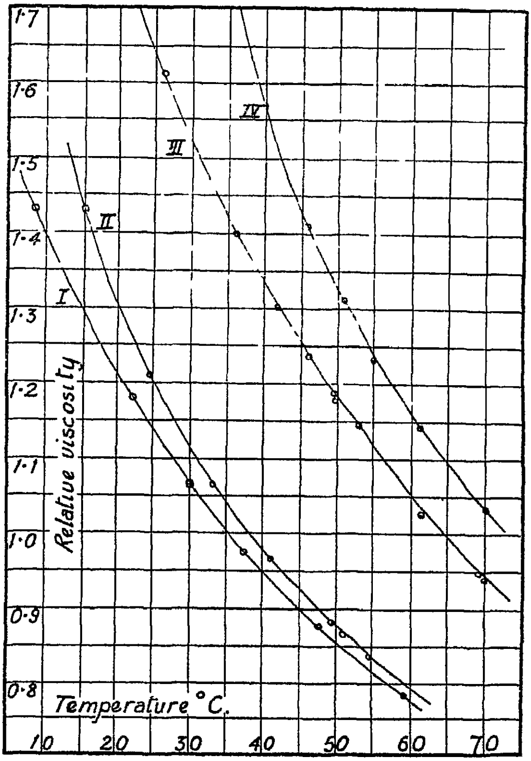


Fig. 2.—Viscosity-temperature curves for ethanolamine oleate dispersions in benzene.

I—0·1 M ethanolamine oleate.
II—0·1 M ethanolamine oleate with 0·35 cc. water added/50 cc.
III—0·3 M ethanolamine oleate.
IV—0·3 M ethanolamine oleate with 1·0 cc. water added/50 cc.

Discussion.

The results obtained in these experiments are readily interpreted in the light of the suggestions made by Schulman and McRoberts and may be taken therefore as confirming these authors' views. Below 50° c. the swelling of the micelle is restricted by the crystalline arrangement of the soap molecules in the lattice. In the neighbourhood of 50° c. thermal agitation overcomes the lateral adhesion of the polar ethanolamine groups, causing the crystalline arrangement of the molecules in the micelle to break down with the formation of a liquid interfacial film which can expand to form droplets. The absence of discontinuities in the viscosity-temperature curves in the region of 50° c. provides satisfactory evidence that the change which occurs in these systems around that temperature is due, not to any change in aggregation of the soap, but to a change in the micelle itself. The only change to be expected on this basis is the very slight one due to the increased volume fraction of the colloid consequent upon the change of state of the micelle at the critical temperature.

It is important to note that the temperature at which the sudden increase in water absorption occurs in these ethanolamine oleate dispersions is not the Krafft point of the soap, which corresponds roughly to the temperature at which the paraffin chains melt, but is considerably lower. The critical temperature in these systems corresponds rather to the point where the last vestige of order in the micelle disappears or, in other words, to the temperature at which the soap in the micelle melts. The absence of a similar temperature effect in the experiments of Schulman and McRoberts is due, no doubt, to the fact that sodium oleate, although well above the Krafft point at 20° c. maintains some order in its lattice to a considerably higher temperature than ethanolamine oleate. The temperature for its transition to a freely flowing liquid is given by Lawrence [6] as 245° c. In view of this requirement it is clear that relatively few soaps would show the critical behaviour with regard to temperature observed with ethanolamine oleate. None of the sodium soaps examined by Lawrence [6] have melting points below 200° c. The reason why ethanolamine oleate melts at the lower temperature and behaves in the manner described is, no doubt, due to the greater difficulty which the ethanolamine groups have, as compared with the sodium ions, in packing into the crystal lattice. On the other hand, search might reveal a soap, probably organic, which would show the critical effect around room temperature. Dispersions of such a soap might have some practical applications.

Summary.

The amount of water absorbed by benzene solutions of ethanolamine oleate increases greatly above a critical temperature. This change is due not to any alteration in the degree of aggregation of the soap at this temperature but to a change in state of the micelle. Above the critical temperature liquefaction of the soap occurs, permitting expansion of the swollen micelles to droplets of visible dimensions.

The Sir Donald Currie Laboratories,
The Queen's University,
Belfast.

[6] *Trans. Faraday Soc.*, 1938, 34, 666.

STUDIES OF ALUMINIUM SOAPS. VII. ALUMINIUM SOAPS IN HYDROCARBONS. THE GELS AND JELLIES AND TRANSFORMATIONS BETWEEN THEM.*

By James W. McBain, Karol J. Mysels † and Gerould H. Smith.‡

Received 18th July, 1946.

An aluminium soap introduced into a hydrocarbon may behave in a variety of ways. Its behaviour depends on the nature of the soap, the

---

* Study conducted under contract OEMsr-1057 between Stanford University and the Office of Emergency Management, recommended by Division 11.3 of the National Defense Research Council, and supervised by Professor J. W. McBain.

† Present address: Department of Chemistry, The University of Southern California, Los Angeles 7, California, U.S.A.

‡ Present address: Union Oil Company of California, Oleum.

nature of the hydrocarbon solvent, the temperature, and often upon the traces of impurities present in either the soap or the hydrocarbon. Aluminium soaps are always wetted by hydrocarbon solvents ; but they may remain practically inert, or swell to various degrees, or dissolve completely. The systems thickened by swelling aluminium soap may be homogeneous, elastic and stringy ; or short, fracturing on handling, and syneretic ; or they may result in several stable layers of different composition and properties. These systems may be clear or turbid, and after heating they may return either to substantially the same state on cooling or to one possessing quite different characteristics.

This complexity of behaviour is reduced in systems composed of a pure aluminium soap and a pure hydrocarbon, which are the subject of the present paper. It is only after such simpler systems are understood that additional components may be introduced, under carefully controlled conditions, to determine their effects. Then one may explain the behaviour of more complex systems, which are of considerable practical importance, such as lubricating greases, paints, etc.

Aluminium soaps in various hydrocarbons, pure and mixtures, have been studied and the results reported by a number of investigators including Markowicz,[1] McBain and McClatchie,[2] Fischer and Hooper,[3] De Fano,[4] and Ostwald and Riedel.[5] In general, the studies of these investigators were limited or their results empirical because of the impurity of soaps and solvents employed.

The current study was made using the purest aluminium di-soaps available. These were prepared by metathesis of potassium soaps with aluminium chloride,[6] followed by washing, drying and, finally, extraction to remove excess or free fatty fatty acid. The solvents were the purest obtainable commercially from Eastman Kodak Company, Carbon and Carbide Chemical Co., and others. In addition to the use of pure constituents, care was taken to exclude both moisture and air from the systems studied by drying and sealing them in evacuated glass tubes whenever possible. The study was implemented by direct observation of the soap and hydrocarbon systems in sealed tubes or weighing bottles, as described in this paper, and also by the use of X-ray diffraction, and of the McBain-Bakr sorption balance,[7, 8] osmometry,[9] and viscosimetry.

The Phases Observed in the Aluminium Soap and Hydrocarbon Systems.

An aluminium soap placed in a hydrocarbon and allowed to come to equilibrium may form a jelly, a sol, or any one of a number of gel forms. In general it has been found that there is no sharp boundary between the jelly and the sol but only a continuous transformation.[10, 11, 12, 13] On the other hand, the differentiation between a gel and a jelly or sol will be shown to be the salient consideration concerning the characteristics of these thickened hydrocarbon systems. Therefore, the following definitions elucidate this very important difference.

---

[1] Markowicz, *Farben-Z.*, 1928, **34**, 326, 414, 503.

[2] McBain, and McClatchie, *J. Physic. Chem.*, 1932, **36**, 2567.

[3] Fischer and Hooper, *Kolloid Z.*, 1930, **51**, 39.

[4] De Fano, *Ind. Ol. min. grassi*, 1929, **9**, 105.

[5] Ostwald and Riedel, *Kolloid Z.*, 1934, **69**, 185 ; 1935, **70**, 67.

[6] Smith, McGee, Pomeroy and Mysels, *J. Amer. Chem. Soc.* (communicated after submission of this paper).

[7] McBain and Bakr, *J. Amer. Chem. Soc.*, 1926, **48**, 690.

[8] Zentner, *J. Colloid Sci.* (communicated).

[9] Working (not yet communicated).

[10] Laing and McBain, *J. Chem. Soc.*, 1920, **117**, 1506.

[11] McBain, *Union Internationale de la Chimie Pure et Applique*, June 19, 1923 ; *J. Soc. Chem. Ind.*, 1923, **42**, 615 ; *The Theory and Application of Colloidal Behaviour* (R. H. Bogue, ed.), (McGraw-Hill Book Company, New York, 1923) ; p. 410, *Chim. et Ind.*, 1924, **11**, 1.

[12] McBain, *Colloid Symp. Monog.*, 1926, **4**, 4.

[13] McBain, *Kolloid Z.*, 1926, **40**, 4.

1. **Gels.**—Systems of two or more components involving a crystalline phase of swollen soap, frequently intermeshing visible or microscopic regions of a second phase which may be liquid-crystalline, a jelly, or a sol. Varieties of gels differ according to whether the crystalline soap is compacted or dispersed, coherent or non-coherent, together with all intermediate stages. Gels are usually opaque, but may be translucent. However, in rare cases, as a result of a suitable combination of refractive indices of the two phases, they can be transparent.

2. **Jellies and Sols.**—Single-phase, transparent, isotropic, colloidal systems, in which there is a continuous gradation from fluid sol to stiff and elastic jelly. Neither of these contains merely swollen crystalline soap.

Four properties have been used for distinguishing between gel and jelly.

Appearance.—The jelly is perfectly transparent and shows no opalescence. It forms no sharp boundary with sol and will swell without limit in added solvent. Gels vary from nearly transparent and faintly opalescent to white and opaque; the opalescence may still be visible at dilutions of about 0·05 % of swollen soap. When the gel is in contact with an excess of sol the separation between the two is easily visible and sharply defined.

Mechanical properties.—The jelly exhibits a continuous, homogeneous, rigid, elastic structure with no syneresis, while the swollen soap is crumbly and clearly discontinuous unless compacted by compression. The jelly heals if sheared, regaining its original structure. The swollen soap, as gel, does not reunite when fractured but remains in discrete pieces and usually exhibits syneresis, if not already surrounded by sol. Bubbles rise freely through a soft jelly and appear spherical, demonstrating its homogeneity. They are often misshapen in a gel, as a result of discontinuous areas.

X-ray diffraction of the residue after evaporation.—A jelly, sol or gel may be frozen by rapid cooling and then exposed to a hard vacuum until all of the frozen solvent sublimes. The soaps thus recovered generally return rapidly to their original state when placed again in the solvent. The residue from a jelly returns to a jelly and that from a gel returns to a gel. This shows that the character of the soap in the gel or jelly is retained on removing the frozen solvent and thus any difference could be detected by suitable tests. These recovered soaps may be distinguished clearly by their X-ray diffraction patterns.[14, 15, 16] The soaps exhibit varying degrees of crystallinity: that recovered from jelly being almost amorphous, while the residue from the gel is not much less crystalline than the original sample. They are, therefore, of a different physical form [17] although they have the same melting-points and may be transformed into each other by proper treatment with hydrocarbon solvents.

X-ray diffraction of the system itself.—This method [16] is not sensitive and can be applied only to concentrated systems. The gel shows all the lines of the original crystalline soap, although with reduced intensity. In the jelly, if retaining some gel fragments, these lines become blurred; otherwise they disappear completely and are replaced by faint spacings and halos.

[14] Ross and McBain, *Oil and Soap*, 1946, **23**, 214.

[15] Marsden, Mysels, Smith and Ross, *J. Amer. Chem. Soc.* (communicated since submission of this paper)

[16] Marsden, Mysels and Smith, *J. Colloid Sci.* (communicated since submission of this paper).

[17] Mysels and McBain, *J. Amer. Chem. Soc.* (communicated since submission of this paper).

These four methods generally permit detection of small amounts of soap in the presence of large amounts of jelly. The detection of small amounts of sol or jelly in swollen soap on the other hand is very difficult because of the intimate intermeshing of the two. Therefore in the following the term two-phase will be used only for those systems in which a clear segregation of a sol or jelly from a gel occurred.

The Transformation between Gel and Jelly or Sol.

A small amount (1-5 %) of aluminium dilaurate, $AlOHL_2$, placed at room temperature in an excess of cyclohexane swells only to an opalescent gel surrounded by a very dilute, very fluid sol. On heating the gel, it swells further, while the concentration of the sol increases. Either before or after the gel has filled the whole liquid, the opalescence disappears and a jelly or sol then exists. The actual temperature of transformation of the gel to the jelly can be determined within 1° or 2° C. when the system is visually examined using proper side illumination against a dark background. As previously stated, this method of examination is believed to be sensitive to at least 0·05 % of soap in the gel phase. As the jelly develops the whole thickens and becomes elastic.

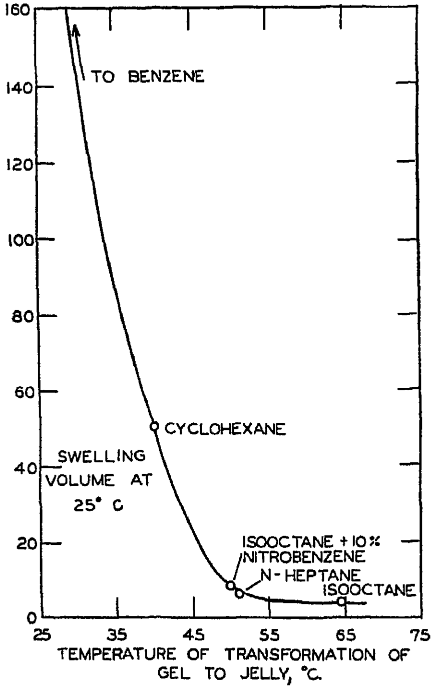


FIG. 1.—Swelling volumes VS. temperatures of transformation of aluminum in hydrocarbons.

This phase transition is subject to modification by a number of factors. The physical form of the aluminium soap, the particular hydrocarbon, the particular aluminium soap, and the presence of impurities or additives such as water, fatty acid, glycols, etc., are all important in their effect on the temperature of this transformation. The main experimental difficulty relating to the temperature of the change of gel to jelly is that of producing uniform distribution of the soap throughout different parts of the system. The soap must be dispersed to allow establishment of equilibrium between the two phases. Fortunately, in the usual range of soap concentration (1-10 %), the soap content has only a small influence upon the temperature of the complete transformation and reliable results can easily be obtained.

Effect of the Physical Form of the Soap.—The fact that aluminium soaps can exist in several physically different forms has [17] been referred to earlier. These forms were found to differ in the extent to which they swelled in a given hydrocarbon, but they also differ as to the temperature at which they pass into the jelly or sol phase. Thus the " amorphous " form of aluminium dilaurate forms a jelly with cyclohexane at room temperature and shows unlimited swelling while the " original," more crystalline, form swells about 50 volumes and passes to a jelly only upon heating to 40° C. The crystallised form imbibes only about 20 weight per cent. without any visible swelling and then does not become a jelly until 48° c.

Effect of Solvent.—The nature of the solvent has a marked effect on the swelling of an aluminium soap and the temperature at which gel is transformed into jelly or sol. The behaviour of systems containing 2·5 % of aluminium dilaurate in various hydrocarbons is shown in Fig. 1. The volume to which the soap swelled at 25° c. is plotted against the corresponding temperature of the change to jelly. The points fell on a smooth line, showing that the temperature of transformation decreases regularly as the swelling increases at a given temperature.

TABLE I.—TEMPERATURES OF TRANSFORMATION AND SWELLING VOLUMES OF ANHYDROUS ALUMINIUM DILAURATE IN VARIOUS HYDROCARBON SOLVENTS.

Solvent.	Swelling at 25° C.		Temp. of Trans., °C.
	Computed Conc. of Gel, %.	Swelling Volume.*	
iso-Octane	24·5	4	64
n-Heptane	18·5	6	50
Cyclohexane	2·0	50	40
Benzene .	—	—	25

* Cc. of settled gel per gram of dry soap.

Effect of a Modifier.—The presence of impurities or addition of modifiers usually has been found to result in a lowered temperature for the change from gel to jelly. For example, the use of 10 % of nitrobenzene in cyclohexane, thickened by 2·5 % of aluminium dilaurate, caused a change from a two-phase gel-sol system to a flowing jelly. The same proportion of nitrobenzene in *iso*-octane lowered the temperature of transformation to jelly from 64 to 50° c. when the system was thickened with 2·5 % of aluminium dilaurate. A point was placed in Fig. 1 for this latter modified system and it seemed to fall on the line drawn through the points for the unmodified systems. Thus it suggested that some modifiers might act by modifying the solvent properties of the hydrocarbon.

The Reverse Transformation from Jelly to Gel.

The reversion from the jelly or sol to the gel on cooling with precipitation of swollen crystalline soap is usually much slower than the change occurring on heating. It does occur in some systems more readily than in others and the nature of both the soap and the solvent seems to be particularly important. In systems such as aluminium dilaurate in cyclohexane, the reversion is slow but occurs if the system is not heated much above the temperature of transformation. However, if it has been heated 50° to 100° c. above that temperature then the jelly may remain clear and stable for months after cooling. Apparently, seeds or crystalline nuclei that promote the reversion may be present in the clear jellies but dissipate upon extended or high heating.

Syneresis always accompanies the reversion of a jelly to form a gel in dilute systems. The gelled portions of the system also become opalescent or semi-opalescent. Upon reheating, the system is again transformed into a jelly and thus the process is reversible. In some cases an increase of 2°-10° c. in the temperature required for the transformation of gel to jelly has been observed on reheating systems that have reverted to the gel form. This suggests that the soap precipitated during reversion is of a more crystalline form than the original sample. This is reasonable as the soap recrystallised slowly from the solvent while during the process of aqueous precipitation the velocity of aggregation certainly was more important than any velocity of orientation. Therefore, in comparing the effects of different solvents, the same batch of soap was used and the transformation occurring with the first heating recorded.

Comparison of the Transformation of Different Aluminium Soaps from Gel to Jelly or Sol.

The three aluminium di-soaps studied, the dilaurate, distearate and dioleate (impure !) all form jellies at about the same temperature in cyclohexane although they differ markedly in their swelling volumes at 25° c. The stearate swells markedly less, the oleate much more than the dilaurate. However, the stability of these various jellies on cooling is quite different. The aluminium dilaurate jelly may remain metastable for quite some time while the distearate jelly changes to a white, opaque, mushy gel very soon after cooling and the dioleate reverts to clear discrete lumps of gel dispersed in a large excess of clear sol. Data on these samples are shown in Table II. A sample of aluminium dinaphthenate was tested along with the di-soaps of the fatty acids. This soap did not form a homogeneous jelly even when the whole system had been filled, but formed a blobby, semi-fluid gel at 85° c. On cooling, the sample remained thickened in this way for a few days then slowly the amount of clear fluid sol increased until only about one-fourth of the system remained thickened as an oily layer.

A sample of aluminium monolaurate [18] was heated in cyclohexane in a sealed glass tube. It showed very little tendency to swell until 70°-90° c. Then it swelled but did not form a thickened system analogous to a jelly at 98° c., even when 7·7 % of aluminium monolaurate was used. This was five-fold the dilaurate that was necessary to form a rigid homogeneous jelly at 40° c. in cyclohexane. Thus the aluminium monolaurate is very inferior to the dilaurate as a thickener of hydrocarbon solvents, at least at lower temperatures. Data relative to the system of aluminium monolaurate in cyclohexane are included in Table II.

TABLE II.—TEMPERATURES OF TRANSFORMATION OF ALUMINUM SOAP GELS TO JELLIES.

Aluminum Soap.	Temp. of Trans., °c.	Conc. of Soap.	Reversion.
Dilaurate . .	40	2·5	May be very slow
Distearate .	42-44	1·5	Rapid
Dioleate . .	42-44	1·5	Rapid
Dinaphthenate *	>85	1·4	Slow
Monolaurate	>98	7·7	Rapid

* Average mol. wt. of acids = 210.

Stability of Jelly Phase.

The jelly form is truly stable at temperatures intermediate between those at which the gel and sol are stable. It is possible that the jelly form may be stable only above a certain concentration in cyclohexane, since it has not been observed in any system containing less than about 1·5 % of aluminium dilaurate. The gel and sol seem to be the only stable forms in these systems. However, at higher concentrations, 2 to 10 % of aluminium dilaurate, the system passes spontaneously to jelly from the gel on heating to 50°-55° c. It is reformed below 100°-150° c. on cooling the sol. Thus the jelly is stable with respect to both gel and sol in this region. The existence of stable gels and sols has been demonstrated previously with other colloidal systems but this is a clear demonstration of the true thermodynamic stability of a jelly repeatedly asserted many years ago by McBain.[10, 11, 12, 13]

From a practical point of view, a jelly may often be unstable and tend to age into a sol or gel, and then either liquefy or " set up " with syneresis.

[18] McGee (not yet communicated).

However, it should be emphasised that there can exist a set of conditions under which a jelly is stable indefinitely. These conditions depend on the particular system but are subject to some control by the use of modifiers and additives.

Phase Diagrams.

A phase diagram should represent the regions defining conditions of temperature, pressure and composition within which various phases of a system exist. It is usual to represent on such diagrams the conditions of thermodynamic equilibria under which the most stable phases in each region are delimited, and the heterogeneous region between such phases in equilibria are also shown. Only very incomplete equilibrium phase diagrams can be drawn yet for aluminium soaps and hydrocarbons.

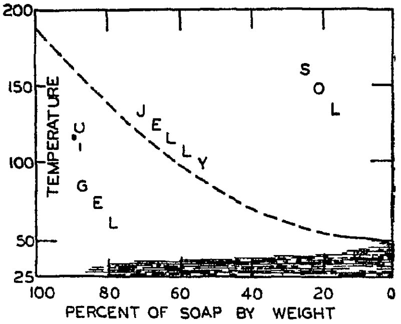


FIG. 2.—Outline of phase diagram of the system "crystallized" Al[OH]L_2—cyclohexane.

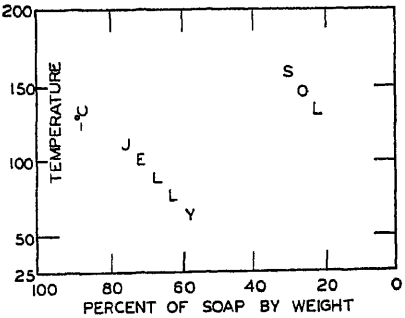


FIG. 3.—Outline of Phase diagram of the system "recovered from jelly" Al[OH]L_2—cyclohexane.

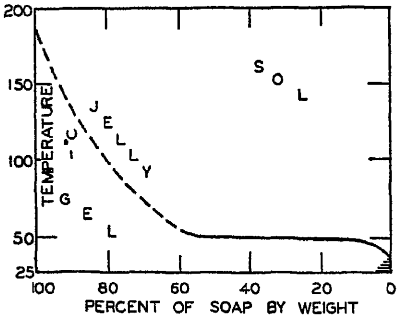


FIG. 4.—Outline of phase diagram of the system Al[OH]L_2—cyclohexane

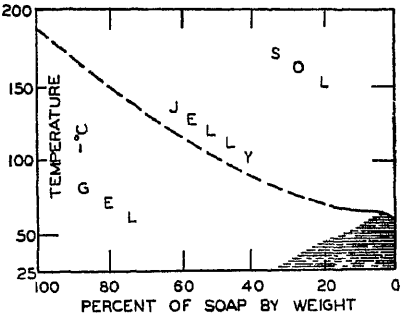


FIG. 5.—Outline of phase diagram of the system Al[OH]L_2—isooctane.

Partial phase diagrams or maps have been drawn to include metastable systems and to illustrate some differences between the systems which have been described in this paper. Figures 2, 3 and 4 show the difference in appearance of the systems of various physically different forms of aluminium dilaurate in cyclohexane. A comparison between cyclohexane and *iso*-octane has been made in Fig. 4 and 5.

Summary.

1. Gels and jelly systems bridge the gap between dry crystalline aluminium soaps and their sols in hydrocarbons. The gels are discontinuous systems of crystalline soap, more or less swollen, frequently

intermeshed with sol or jelly The jelly is a one-phase, homogeneous, transparent isotropic elastic system

2 These forms have been studied by the visual appearance, mechanical behaviour and X ray diffraction of the systems themselves as well as the residues remaining when the systems were frozen and the solvent removed by sublimation

3 The temperature of the transformation of a gel to a jelly or sol was found to depend on the physical form of the soap the nature of the solvent the nature of the soap and the presence of impurities or additives This transformation was found to be reversed upon heating and cooling

4 The jelly is clearly demonstrated to exist as a thermodynamically stable phase within a region between those in which the sol and the gel are the stable forms

5 Diagrams including unstable phases are drawn to illustrate the differences of behaviour which could be attributed to the physical form of the soap and to the hydrocarbon solvent

Department of Chemistry,
Stanford University,
California

SWELLING AND MOLECULAR ORGANISATION IN COLLOIDAL ELECTROLYTES.

By D G Dervichian

Received 7th August 1946

In general studies on swelling are concerned with high polymers, or in any case with substances with very large molecules Despite the imbibition the colloid retains some rigidity and the swollen phase is often a gel In this paper we shall consider concentrated phases of pure or associated lipids, phospho lipids, sterids etc , which are in equilibrium with an excess of water or a more dilute phase These concentrated phases are more plastic than rigid They are definitely distinct from gels which, however can be obtained in other conditions, with some of the same substances Stress will be laid especially on the characteristic property that swollen phases have of reaching equilibrium limit in presence of the penetrating liquid The substances which are dealt with have relatively small molecules of known structure, and some of their aqueous solutions have already been analysed by X ray diffraction methods It is therefore possible to derive some precise information on the distribution and role of water in the structure of the swollen phase

It may be helpful at the outset to describe briefly the different forms of swelling given by these substances[1] myelinic figures anisotropic droplets coacervates and peptisation By far the most typical and best known is the swelling of lecithin in water[2] [3] [4] When examined under the microscope, between slide and cover slip, a particle of lecithin in contact with water swells on the periphery then cylindrical excrescences appear which grow either in simple or twisted fashion, but each keeping a constant diameter These forms are called *myelinic figures,* because the myelin of the nerve sheath or alcoholic brain extract behave similarly (Fig 1 and 2 (Plate III)) The submicroscopic structure of these figures

---

[1] Dervichian and Magnant, *Compt rend biol* 1946 **140**, 94

[2] Virchow *Virchow s Arch path Anat Phys* , 1854, **6**, 562

[3] White *The Medic Chron* , 1908 47

[4] Nageotte, *Morphologie des Gels Lipoides*, (Hermann et Cie, Ed Paris, 1937)

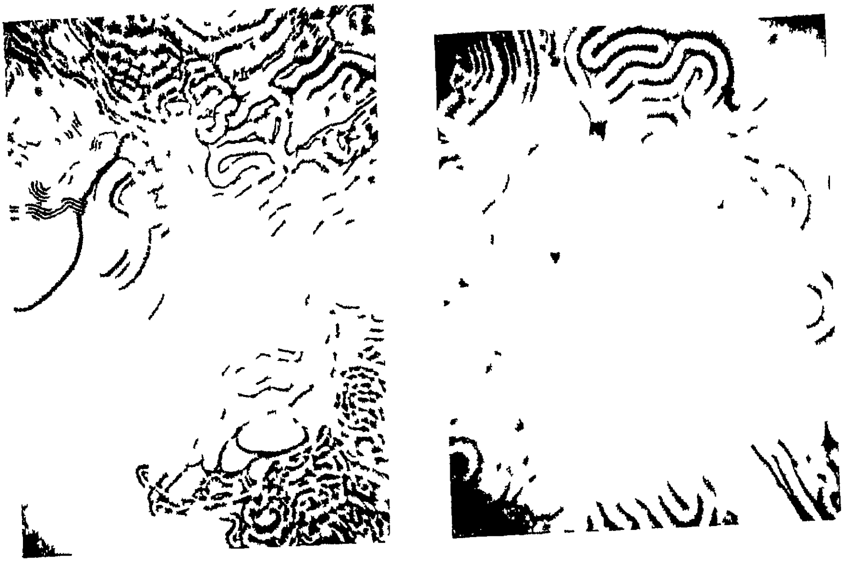


FIG. 1 and 2.—Myelinic forms obtained with stearylcholine hydrochloride + cholesterol in water

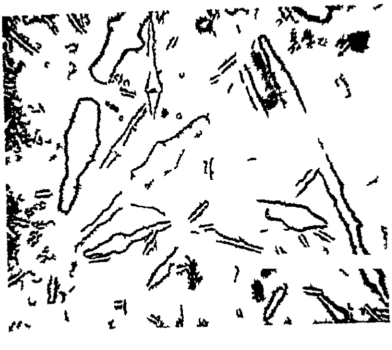


FIG. 3.— Batonnets obtained with sodium laurylsulphate + cholesterol in water

FIG 4.—Tactoids given by tripalmitin + bile acid salts in water

[*To face page* 180

will be described later. We shall anticipate by saying that they are liquid crystals or "paracrystals",[5] fluid in the direction of their growth, and resistant in the direction perpendicular to the axis of the cylinders. These cylinders are birefringent, and do not coalesce when they come into contact. They contain a considerable proportion of water, but remain in equilibrium with the surrounding water without dispersing in it.

A solid particle of a mixture of cholesterol and sodium-laurylsulphate swells in water and disaggregates into droplets having more or less elongated shapes, with axial symmetry. They present all the characters of the *batonnets* described by Friedel[6]: they are birefringent, and show characteristic bulges (Fig. 3). These anisotropic droplets when they come into contact coalesce and form new particles of the same type. In some cases spindle-shaped droplets are formed (*tactoids*) similar to the particles observed in virus protein preparations (Fig. 4). Here also the paracrystalline phase contains a great amount of water, but is, nevertheless, in equilibrium with the surrounding water.

Spherical and isotropic drops can be obtained with other mixtures containing long chain alcohols associated with bile-acid salts. These are the *coacervates* which have been studied principally by Bungenberg de Jong and his collaborators.[7] In the phenomenon of coacervation, a mixture of two colloidal solutions separates into two phases, one of which may contain the dissolved colloids at a concentration of 10 % while the other, under certain conditions, is nearly pure water. The interfacial tension between the two phases is so small that the slightest shake scatters the phase called coacervate into tiny isotropic droplets remaining in equilibrium in the surrounding water and coalescing again when they happen to meet. This scattered phase has been called *micro-coacervate* by Bungenberg de Jong.

It should be emphasised that these different forms of equilibrium are obtained always by associating an insoluble molecule (cholesterol, triglyceride, long chain alcohol, etc.), with a soluble substance (alkyl sulphates, bile-acid salts, digitonin, etc.). In view of the general theory of the phenomenon, it is interesting to mention the extreme case of complete dispersion. Thus, by adding colchicin (which is soluble) to cholesteryl phosphate (which is insoluble), complete peptisation is obtained. This is due to formation of a complex in the micelle which exists in solution.

It should be noted that most of the substances investigated have, following G. S. Hartley's terminology, amphipathic molecules. Soaps represent the most typical of these substances, and it is of interest to describe the condensed phases which are met with in their solutions. The numerous systematic investigations of J. W. McBain and his school[8] furnish precise data on the conditions of existence and equilibrium of different phases in aqueous soap solutions. Furthermore, X-ray analysis of the same solutions carried out in the course of the last few years[9, 10, 11] provides precise information on their structure and on the distribution of the water.[12] Most of the investigators have been particularly impressed by the presence of ordering in these solutions which gives rise to very large

[5] Rinne, *Trans. Faraday Soc.*, 1933, **29**, 1016.

[6] Friedel, *Ann. Physique* (9), 1922, **18**, 328.

[7] Bungenberg de Jong and Kruyt, *Proc. acad. sci. Amsterdam*, 1928, **32**, 849; *Kolloid Z.*, 1930, **50**, 30. Holleman, Bungenberg de Jong and Dekker, *Kolloid Beihefte*, 1934, **43**, 143. Bungenberg de Jong and Saubert, *Ned. Akad. Wetensch.*, 1942, **45**, 401. Bungenberg de Jong and Meer, *ibid.*, 498.

[8] McBain and Langdon, *J. Chem. Soc.*, 1925, **127**, 852. McBain, Lazarus and Pitter, *Z. physic. Chem. A.*, 1930, **147**, 87. McBain, Vold and Vold, *J. Amer. Chem. Soc.*, 1938, **60**, 1866.

[9] Hess, Kiessig and Philippoff, *Naturwiss.*, 1938, **26**, 184.

[10] Stauff, *Kolloid Z.*, 1939, **89**, 224.

[11] Kiessig and Philippoff, *Naturwiss.*, 1939, **27**, 593.

[12] For a systematic description of the structure of soap solutions, see Dervichian and Lachampt, *Bull. Soc. Chim.* 1945, **12**, 189.

spacings at relatively dilute concentrations. These rather dilute solutions are certainly interesting in themselves, but we shall be concerned here with the structure of condensed phases which have a more direct relation with the swelling phenomenon.

Anisotropic Phases in Concentrated Soap Solutions.—One point should be kept constantly in mind in order to understand the behaviour of soaps in the presence of water. Colloidal solutions of soap appear only above a temperature which is nearly constant and has a definite value for each kind of soap. This temperature is called the Krafft point, and depends on the nature of the fatty acid and of the cation.[12a] All the particularities described below occur above the Krafft point. The temperature should be above 70° c. for sodium stearate or 50° c. for potassium stearate, while sodium oleate gives solutions already at room temperature. Below the Krafft point suspensions of micro-crystals or metastable gels are obtained. By X-ray examination of soap solutions at different concentrations above the Krafft point Hess,[9] Stauff,[10] and Kiessig and Philippoff [11] have been able to show that the long spacing, corresponding to the width of the bi-molecular layers, increases with dilution. This study of soap solutions has since been repeated by other workers [13, 14] without bringing to light much new fundamental information. The most obvious interpretation was given from the beginning. Water penetrating between the polar groups increases the long spacing. Moreover, while the anhydrous crystal shows two principal side spacings as a consequence of the arrangement of the molecules in a monoclinic lattice, only one side spacing, intermediate between those of the anhydrous crystal, is given by the solutions. It has been concluded that the molecular chains are free to rotate in the micelle, and that their arrangement corresponds to hexagonal close packing, the state being similar to the smectic state.

To evaluate the thickness of the water intercalated between the soap layers, one can obviously subtract from the long spacing measured for a given concentration the thickness corresponding to the soap alone. There is no reason to adopt for this thickness of the soap layer either the calculated length of a fully extended double molecule * or the value of the long spacing measured in the anhydrous soap crystal in which the molecules are tilted in a monoclinic lattice. It is much more reasonable to take for the thickness of the soap double layer the value determined experimentally on the anhydrous soap in a state which is similar, if not identical, to the state which exists in the dissolved micelle. Thus no assumption is made about the arrangement of the molecules in the layers. Anhydrous sodium stearate at temperatures ranging between room temperature and 301° c. (melting point = 288°) has been studied by X-ray diffraction methods by de Breteville and McBain.[15] As the temperature increases, the anhydrous soap passes through different phases before reaching the liquid state. In the neighbourhood of 170° c., there is a transition from the monoclinic to the hexagonal close packed arrangement, the molecules exhibiting cylindrical symmetry. At the same time, the long spacing decreases gradually, reaching a value of 32 A. for the liquid state when the calculated length of a fully extended double molecule is 52 A. In the course of this evolution, and before reaching the liquid state, the soap passes by the *neat soap* phase which is analogous to the phase which occurs in concentrated solutions. It is therefore, reasonable, to adopt for the thickness of the soap double layer the long spacing measured in the anhydrous neat state, i.e. 33 A. for sodium stearate.

---

* If the calculated length of a fully extended double molecule is taken as the thickness of the soap layer, paradoxical results are obtained. In fact, the long spacing measured in sufficiently concentrated solutions is shorter than this calculated length, although water is intercalated between the layers [13, 14].

[12a] Démarcq and Dervichian, *Bull. Soc. Chim.*, 1945, **12**, 939.

[13] Ross and McBain, *J. Amer. Chem. Soc.*, 1946, **68**, 296.

[14] Harkins, Mattoon and Corrin, *ibid.*, 220.

[15] de Bretteville and McBain, *J. Chem. Physics*, 1943, **11**, 426.

Unfortunately, Stauff's [10] measurements were made on sodium laurate without taking any account of the existence of the different phases. Nevertheless, the long spacing corresponding to the paracrystalline phases can be evaluated by interpolation. Thus *ca.* 32 A. would correspond to the *neat* phase and *ca.* 37 A. to the *middle* phase. In order to evaluate the thickness of the soap double-layer alone, it may be assumed that the shortening of the chains in the hexagonal packing measured for sodium stearate is proportional to the chain length. This leads to the value of 25 A. for the double layer of sodium laurate, instead of the calculated 37 A. for the fully extended double molecule. Thus the thickness due to the intercalated water would be approximately 7 A. in the *neat* phase, and approximately 11 A. in the *middle* phase. Taking 3 A. as the dimension of a water molecule, this corresponds to 2 and 4 layers respectively of water molecules. As the middle phase immediately precedes the isotropic solution in order of increasing dilutions, one can say that the paracrystalline state disappears as soon as more than 4 layers of water penetrate between the soap layers.

Knowing from X-ray data the volume of water located between the micelles and, on the other hand, the total quantity of water in the soap, a simple calculation enables us to determine the volume of water placed laterally around the micelles for each concentration.[16] This calculation is obviously very crude, but gives an idea of the order of magnitude of the mean lateral distance between micelles. For example, in the 30 % soap phase, the average lateral distance between the micelles is equal to half the average lateral dimension of the micelles, whatever this dimension may be. In the 70 % phase, the lateral interval represents only one-sixth of the lateral dimension. It has been said that paracrystalline phases of the smectic type, as are these concentrated soap phases, are crystalline only in the direction perpendicular to the parallel layers. From this point of view, the water between the layers can be considered as water of crystallisation taking part in the structure of the soap. From the swelling point of view, it is also the imbibition water. On the other hand, there is no favoured direction in the isotropic solution ; here the soap is dispersed in the water. This equilibrium between paracrystal and solution can be compared, for example, with that existing between a crystal of Na_2CO_3, $10H_2O$, and its saturated solution at 30°, the crystals contain 37 % of anhydrous Na_2CO_3 and the solution 28 %. For any concentration between 28 and 37 % the mixture separates into two phases. This interval is analogous to the " demixtion zones " in McBain's diagrams. With $CaSO_4$, $2H_2O$ the interval is much wider, the concentrations being respectively 79 % and 0·24 % of anhydrous $CaSO_4$ in the crystal and in the solution. The case of acid soaps is illustrated by this last example.

Swelling of Acid Soaps and Phospholipids—Acid soaps have been known for a long time, and were described by Chevreul. A method of preparing potassium dioleate has been given by McBain and Stewart.[17] The behaviour of fatty acids as di-acids cannot be explained on chemical grounds. There are strong reasons for believing that acid soap results from the co-ordination of acid and neutralised molecules in the crystalline lattice in a similar way to the arrangement of the mixtures which are described below.

Acid soaps are very poorly soluble, but swell in water giving paracrystalline phases. Compared with the neutral soaps the " demixtion zone " is considerably widened in the case of acid soaps. In the particular case of acid sodium caprylate described by Ekwal,[18] an anisotropic phase is formed which contains 30 to 40 % water, while certainly less than 0·5 % soap is dissolved in the excess of water. Although the association of the insoluble free acid with the neutral soap considerably reduces the solubility of soap in water, water can still enter into the crystalline structure of the soap giving what we can call the swollen phase.

---

[16] Dervichian, *J. Chim. Physique*, 1941, 38, 63.

[17] McBain and Stewart, *J. Chem. Soc.*, 1927, 130, 1392.

[18] Ekwal, *Kolloid Z.*, 1941, 97, 71.

Schmitt, Palmer and Bear [19, 20, 21] have applied X-ray methods to the study of phospholipids dispersed in water. Measurements have been made on what they call "emulsions" of 1 part of phospholipids in 2 of water. Actually, as far as lecithin is concerned, this is neither an emulsion nor a dispersion, but consists of fragments of myelinic figures.[22] One of the interesting points found by these authors is that lecithin gives only a limited increase of the long spacing. The maximum value is 69 A. It is difficult to evaluate the number of intercalated water layers as different values are found for the long spacing in the dry lecithin: one is 43·4 and the other 62·2 A. The authors point out, however, that heating at 100° C. tends to favour the 62 A. spacing at the expense of the 43 A. spacing. It is difficult to know how far this spacing is related to the thickness of the lecithin double layer in the swollen phase. All we can say is that only a few layers of water are intercalated.

Cephalin, on the contrary, shows a spacing of 123 A. In this case, speaking of a concentration of 1 in 2 of water has a real significance, for cephalin disperses completely in water. The solution is comparable to the isotropic solution of neutral soaps. On the other hand, the swollen phase of lecithin is equivalent to the anisotropic condensed phases of neutral soaps or, better, to the swollen phase of acid soaps.

It is very instructive to compare the molecular structure of lecithin with those of two related substances: lysolecithin and triolein. Triolein, with three hydrocarbon chains and without hydrosoluble groups, is completely insoluble in water. Lysolecithin, on the contrary, disperses in water because the molecule has only one hydrocarbon chain, and besides the free OH group of the glycerol it contains a strongly solubilising phosphorylated amino-alcohol group. Lecithin stands between these two extreme cases, with its two insoluble hydrocarbon chains counterbalancing the hydrosoluble character of the phosphorylated choline group. Therefore, lecithin swells although it does not dissolve readily in water.

The Notion of Hindered Solubility and the Swelling of Paracrystalline Complexes.—What is found naturally in the lecithin molecule can be realised artificially with a mixture. Lysolecithin, associated with cholesterol in equimolecular proportions, swells and gives myelinic figures, as does lecithin alone. The insoluble cholesterol molecule hinders the solubility of the lysolecithin molecule, exactly as free fatty-acid does with the neutral soap. Besides, the same result can be obtained by associating with the cholesterol, sodium cetylsulphate or stearylcholine hydrochloride. The cholesterol itself can just as well be replaced, cetyl alcohol being taken as the insoluble substance and associated with a soluble sterid derivative, digitonin.[22] Thus is introduced the idea of hindered solubility which has been already used to explain the formation of coacervates.[23]

In this way we may expect to prepare a whole graduated series of phases by choosing as components molecules whose solubility or insolubility is more or less marked. As reported above, we have thus been able to obtain a great number of different swollen phases ranging from myelinic figures to complete dispersion.[1]

Molecular Organisation of Swollen Phases.—It has been shown by the X-ray work of Bear, Palmer and Schmitt [20] that in lecithin and similar substances, the water of imbibition is located between the parallel

---

[19] Schmitt and Palmer, *Cold Spring Harbour Symp.*, 1940, **8**, 94.

[20] Bear, Palmer and Schmitt, *J. Cell. Comp. Physiol.*, 1941 **17**, 355.

[21] Palmer and Schmitt, *ibid.*, 385.

[22] Dervichian and Magnant, *Bull. Soc. Chim. Biol.*, 1946, **28**, 419.

[23] Dervichian, *Compt. rend.*, 1944, **219**, 679.

sheets of the lipid molecules. On the other hand, both optical observations [24] and the mechanical properties as examined with a micro-manipulator [25] prove that, in myelinic figures, the cylinders are formed by coaxial cylindrical sheets fitting into each other. We are thus fully justified in assuming that the macroscopical cylindrical sheets consist of a very great number of molecular double layers separated by a fixed number of layers of water and rolled one inside the other.

The definite molecular structure of the *batonnets* of Friedel still remains to be elucidated. There are, however, strong reasons for believing that this structure also is lamellar.[6, 26] Finally, on the basis of very crude assumptions and calculations, the conclusion has been drawn that in *coacervates* the water surrounding the colloidal particles must represent at most three or four molecular layers.[23]

It remains to define how the molecules of two components are associated in the swollen phases obtained by mixing soluble and insoluble constituents. The answer is found from the study of the individual layers by the technique of Schulman and Rideal of monolayer penetration.[27, 28, 29, 30, 31] A parallel study on monolayers has been carried out using all the pairs of substances which gave the different paracrystalline swollen phases.[32] The mixed monolayers thus obtained permit us to fix the stoichiometrical proportions which give rise to the formation of a complex. The same pair of substances has to be mixed in the proportions shown by the monolayer technique in order to obtain swollen phases. Still more convincing is the fact that when a given pair of substances gives complex films for different proportions, only mixtures with the proportions corresponding to *fluid* films can swell in water.[32] The crystallographic meaning of these complexes has been stressed in another paper.[33] They are due to some sort of syncrystallisation in each of the individual layers and the stoichiometrical proportions are the consequences of the regular co-ordination of the two species of molecules side by side in a common two-dimensional lattice. Another fundamental point is that the association is correlated with the swelling, and vice versa.

Interpretation of Swelling.—An interpretation proposed for the coacervation phenomenon [23] can be generalised to explain the formation of paracrystalline phases. It is a question of the balance between the free energy of molecular cohesion * and the free energy of affinity for water. Both these energies depend upon the nature of the paraffinic and ionic lattices: i.e., on the nature of the hydrocarbon tails and of the polar groups of the amphipathic molecules. The pure substance or the mixture is insoluble when cohesion is stronger than affinity, or is soluble in the inverse case. Increasing the solubility can obviously be effected either by increasing the affinity or by decreasing the cohesion. The intermediate swollen state is obtained by adjusting the relative importance of cohesion and affinity. The significance of the Krafft point of soaps—the temperature at which, roughly speaking, the paraffinic chains melt and below which there is neither dispersion nor swelling—thus becomes obvious. It also becomes obvious that a pair of substances which are liable to give a solid monolayer (strong cohesion) cannot give swollen paracrystals. With shorter chains, on the other hand, cohesion is so weak that complete dispersion occurs.

---

* Cohesion as well in the "paraffinic lattice" as in the "ionic lattice", using Hartley's terminology.[34]

[24] Nageotte, *Compt. rend.*, 1927, **185**, 1021.

[25] Browaeys and Dervichian, *Compt. rend. Biol.*, 1946, **140**, 136.

[26] Bragg, *Trans. Faraday Soc.*, 1933, **29**, 1056.

[27] Schulman and Rideal, *Proc. Roy. Soc. B.*, 1937, **122**, 29 and 46.

[28] Schulman and Stenhagen, *ibid.*, 1938, **126**, 356.

[29] Schulman, *Trans. Faraday Soc.*, 1937, **33**, 1116.

[30] Schulman, Stenhagen and Rideal, *Nature*, 1938, **141**, 785.

[31] Joly, *ibid*., 1946, **158**, 26.

[32] Dervichian and Joly, *Bull. Soc. Chim. Biol.*, 1946, **28**, 426.

[33] Dervichian, *ibid.*, 433.

[34] Hartley, *Aqueous Solutions of Paraffin-Chain Salts*, (Hermann et Cie, Ed., Paris, 1936).

It remains to be explained why there is a limit to the amount of water of swelling. It has been shown above that the thickness of water intercalated between the sheets of soap and compatible with the existence of a paracrystalline phase had a maximum corresponding to not more than four layers of molecules. It is well known that this maximum of four layers is also found in swollen montmorillonite. A particular arrangement of the molecules of water in the swollen clays has been proposed by Hendriks and Jefferson [35]. Certain identities in molecular dimensions lead us to adopt the same point of view as them as regards the imbibition water of the colloids which are described in this paper. The water molecules, instead of being tetrahedrally co-ordinated are supposed to be arranged on a plane hexagonal lattice. The consequence is that every alternate water molecule has an H atom emerging upwards out of the plane. In this arrangement the distance between the O atoms of the molecules is given as 3 A., being larger than the 2·8 A. found in the normal tetrahedral structure. Hendriks and Jefferson have shown that with this arrangement there is a close coincidence with the structure of the surface of the clay sheets.

If we calculate the area which can be ascribed to each oxygen atom with an emerging H bond, the value of *ca.* 24 A^2. is found. Now, the side spacing of the long chains in the soap micelles measured by X-ray methods and corresponding to hexagonal close packing (above the Krafft point) is 4·6 A. From this the area per molecule of soap can be calculated and is again *ca.* 24 A^2.* This identity of molecular dimensions allows us to make the assumption that the structure of the adjacent first layer of water is conditioned by the structure of the monolayer in the soap micelle. The action is probably reciprocal. It is reasonable to make the further assumption that the particular hexagonal structure which is imposed on the water between the sheets of soap cannot be maintained with more than four layers of water, and that with more layers the normal tetrahedral co-ordination sets in. In fact the intensity curve of X-ray scattering by liquid water, given by Bernal and Fowler,[37] shows that the position of molecules around a given molecule cannot be defined further than the two or three first shells of neighbours.

Making these assumptions, we have to admit, as a consequence of the experimental facts, that water can no longer enter the crystallographic structure of the swollen phase when it returns to the normal tetrahedral co-ordination. Thus, when the quantity of water is larger than four layers, the solubility of the colloid in the water should be considered. Now, this solubility of the colloid in the water can be of greater or less importance (neutral soap or acid soap or complexes). The classical phenomenon of swelling corresponds to the case where the solubility is very small.

The existence of ionised groups on the surface of the micelles and the presence of an atmosphere of gegenions have been invoked to explain the nature of long range interactions and the organisation in relatively dilute colloidal electrolytes. It is nevertheless difficult to apply this interpretation to the case of penetration of organic non-polar liquids between the paraffinic end groups of the soap molecules inside the micelle.[38, 39] The increase of the long spacing in a soap solution due to the introduction of increasing quantities of benzene intercalated between the non-polar faces of the micelle sheets has been studied by Kiessig and Philippoff.[38] When one examines the experimental points on the figure given by these authors, one notices that, within a certain range of concentration of benzene, the long spacing remains constant, the increase in spacing due to the benzene being 9 A. This suggests that a definite phase has appeared. The 9 A. thickness can only correspond to, at most, three layers of benzene molecules stacked flat. Thus, whether the swelling is due to benzene or to water, there exist definite phases with only a limited number of intercalated layers.

---

* Incidentally this seems to justify the term "mesomorphous" which had been given to a particular condensed fluid state found in monolayers of fatty substances [36] and throws some light on the significance of this characteristic area.

[35] Hendriks and Jefferson, *Amer. Mineralogist*, 1938, **23**, 863.

[36] Dervichian, *J. Chem. Physics*, 1939, **7**, 944.

[37] Bernal and Fowler, *ibid.*, 1933, **1**, 515.

[38] Kiessig and Philippoff, see ref. [11].

[39] Hughes, Sawyer and Vinograd, *J. Chem. Physics*, 1945, **13**, 131.

We should perhaps emphasise that we have not been concerned in this paper with swollen gels, but only with plastic phases. It may be said that the relative fluidity of these phases is due to the fact that they are not composed of high polymers. Yet the soaps themselves can give a gel phase which demonstrates the existence of these two different sorts of structure. The gel, which can only be formed below the Krafft point (cohesion stronger than affinity), appears as a metastable state in the form of long entangled filaments.[40] Water, instead of being bound between the soap layers, is simply held by the entangled mesh of filaments. This structure corresponds to the generally accepted idea of swelling.

Summary.

The conditions of formation and the molecular structure are described of different concentrated phases of colloidal electrolytes in equilibrium with a more dilute phase. By mixing two long chain substances, one soluble and the other insoluble, different types of these equilibria (myelinic figures, anisotropic droplets, coacervates) are obtained at will. A general theory is put forward to explain the limited amount of water which can enter into the swollen phase.

I should like to thank Dr. D. P. Riley, Scientific Adviser, British Council, Paris, for the help he has given me in preparing this paper in English. My thanks are also due to Mr. P. Jeantet and Mlle. C. Magnant for taking photographs of the microscopical preparations.

Service de Chimie physique,
Institut Pasteur,
Paris.

---

[40] Dervichian, *Compt. rend.*, 1943, **217**, 299.

THE COMPOSITION AND SWELLING PROPERTIES OF HAEMOGLOBIN CRYSTALS

BY M. F. PERUTZ.*

Received 7th August, 1946.

It has long been known that protein crystals can undergo swelling and shrinkage, that a large proportion of their volume normally consists of water, and that they are permeable to smaller ions. In recent years their composition has been studied by Adair and Adair,[1] and by McMeekin and Warner,[2] principally with a view to measuring the hydration of proteins.

---

* Imperial Chemical Industries Research Fellow.

[1] Adair and Adair, *Proc. Roy. Soc. B.*, 1936, **120**, 422.

[2] McMeekin and Warner, *J. Amer. Chem. Soc.*, 1942, **64**, 2393.

The present study started as an attempt to determine the exact amounts of liquid in wet hæmoglobin crystals and the electron density in the liquid regions, data which were needed as part of a crystal structure analysis of horse methæmoglobin. It was mainly the difficulties encountered there which led to the study of the ionic equilibria between the interior of protein crystals and their suspension medium; this in turn drew our attention to the interesting swelling properties of the hæmoglobin crystals and raised the problem of the lattice forces which hold them together. The question of protein hydration, though outside the scope of the present discussion, has a direct bearing on the mechanism of swelling in protein crystals and gels, and is therefore treated as part of this investigation.

The lattice forces in protein crystals might be of several kinds, but for the present purpose it will be convenient to distinguish two main types: long range forces such as appear to exist in certain gels, and short range forces of the type more commonly found in crystal structures. Evidence of the former comes principally from studies of tobacco mosaic virus and bentonite gels and, more recently, from experiments with soap and phospholipoid micelles. The interparticle distances in those systems range from tens to several thousands of Ångstrom units, and vary as a function of both hydrogen ion and neutral electrolyte concentration. The nature of the forces which keep particles parallel and equidistant across such great thicknesses of water is not yet clear. If the lattice forces in protein crystals were of this long-range type, the unit cell dimensions should vary with hydrogen ion and electrolyte concentration. On the other hand, if short-range electrostatic forces between hydrated ions predominate, and hydrogen bonds act either directly between protein molecules or through intermediary water molecules, as in certain hydrated crystals, the unit cell dimensions should be independent of the composition of the suspension medium, except when a change in *p*H actually alters the distribution of fixed ionic changes on the surface of the protein molecule.

Before proceeding further, it will be useful to summarise our present knowledge of the structure of horse methæmoglobin crystals.[3, 4] These crystals are monoclinic, belong to the space-group *C*2, and have two molecules in the unit cell. They possess a layer structure, with rigid layers of hæmoglobin molecules extending throughout the crystal parallel to the crystallographic *c* plane, and alternating with layers of liquid of crystallisation whose thickness in the normal wet crystal is 15 A. If the crystals are gradually dried, the thickness of the liquid layers is reduced in steps first to 10 A., then to 6 A., and finally to zero. The layers of hæmoglobin molecules retain a thickness of 36 A. at all degrees of wetness of the crystal; they are rigid and apparently impenetrable to water. Evidence concerning their internal structure need not be described here.

Experimental.

Crystals of horse methæmoglobin were grown either by dialysis of hæmoglobin solutions against concentrated salt solutions or by electrodialysis of highly concentrated hæmoglobin solutions against tap water. In either case well developed crystals of up to 1 mm. diameter were obtained. The effect of different suspension media was studied by first washing them with the new medium on a slide and then leaving them for several days to reach equilibrium in the new medium before taking any measurements. The densities of the wet crystals were determined by flotation in mixtures of xylene and bromobenzene, according to the method of McMeekin and Warner.[2] Adair and Adair's method [1] of flotation in concentrated salt solutions was also employed. The different results given by these two methods can best be visualised by comparing a protein crystal to a wet sponge: McMeekin and Warner float the sponge in a

---

[3] Perutz, *Nature*, 1942, 149, 491.

[4] Boyes-Watson and Perutz, *ibid.*, 1943, 151, 714.

liquid which does not mix with water and thus measure the density of the wet sponge, while Adair and Adair float the sponge in salt solutions; as the salt diffuses freely into the sponge, the property one would expect to be measured by this method is the density of the dry sponge skeleton. Actually, a certain proportion of the water inside protein crystals does not seem to be available to the diffusing salt, so that the observed density is considerably lower than that of the dry protein. In calculating the crystal composition, we also had to use the partial specific volume of the anhydrous protein or its reciprocal, the apparent density. This quantity is the same[1] whether hæmoglobin is present in solution or in the form of a crystalline suspension; it is also the same in the absence of salt or in the presence of M/NaCl. The results obtained with a salt-free methæmoglobin crystal are given in Table I.

TABLE I.

Method of density determination.	Crystal density at 18°C.
1. Flotation of wet crystal in xylene-bromobenzene mixture .	1·160
2. Flotation of wet crystal in concentrated salt solution . .	1·242
3. Flotation of air-dried crystal in xylene-bromobenzene mixture	1·270
4. Apparent density of crystal suspension in water . . .	1·33

The unit cell dimensions of the crystals were determined with the help of oscillation photographs, with an accuracy varying between 1 % and 0·3 %, according to the angle to which the X-ray reflections extend.

Evaluation of Results.

The molecular weight of horse hæmoglobin is known from analysis of its iron content, combined with osmotic pressure measurements, to be 66,700; its molecular volume is obtained by multiplying the weight with the partial specific volume of the anhydrous protein in solution (the justification for this procedure is shown in the previous section). The specific volume of the liquid crystallisation (i.e. the total liquid in the crystal), can now be derived from the equation:

$$v = \frac{V - nMv_p}{VD - nM} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where v = specific volume of the liquid crystallisation, V = unit cell volume in A^3., n = number of hæmoglobin molecules per unit cell M = molecular weight of hæmoglobin in 10^{-24} g., D = crystal density, and v_p = partial specific volume of anhydrous protein in solution.

We now assume that the liquid of crystallisation consists of two distinct components, viz., water which is bound to the protein and therefore not available as solvent to mobile ions ("bound water") and a solution through which mobile ions diffuse freely ("free liquid") and which has the same composition as the suspension medium outside the crystal. It follows that

$$v = xv_{H_2O} + (1 - x)v_m$$

where x is the weight fraction of bound water, v_{H_2O} is the specific volume of water, and v_m that of the suspension medium. The weight fraction of bound water is therefore given by

$$x = \frac{v - v_m}{v_{H_2O} - v_m} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

whence, protein hydration (expressed as grams water per gram protein, w), is

$$w = x\left(\frac{VD - nM}{nM}\right) = \frac{v - v_m}{v_{H_2O} - v_m}\left(\frac{VD}{nM} - 1\right) \quad . \quad . \quad (3)$$

Of the four possible sources of error only those in the crystal density would seriously affect the value of w which should be accurate to ± 0·03, at least.

The Effect of Neutral Electrolyte.

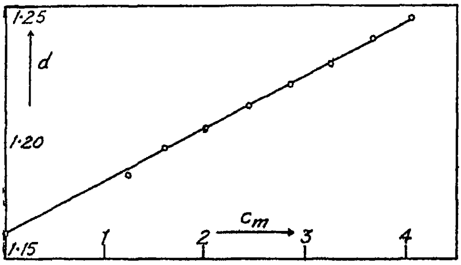


FIG. 1.—Density of wet horse methæmoglobin crystals plotted against ammonium sulphate concentration in the suspension medium c_m = M $(NH_4)_2SO_4/l$.

The unit cell dimensions of horse methæmoglobin crystals were found to be the same whether the crystals had been grown by dialysis against neutralised $(NH_4)_2SO_4$ solution or neutralised K_2HPO_4 solution, or by electrodialysis against tap water. The different methods of growth also had little influence on the distribution of intensities in the X-ray diffraction pattern. In order to test whether the concentration of neutral electrolyte in the suspension medium has any effect on the crystal structure, a crop of crystals was grown by dialysis against $(NH_4)_2SO_4$ solution, and the crystals were then immersed in a series of eight $(NH_4)_2SO_4$ solutions varying in strength between 1·25 and 4·05 molar; this series was prepared by successive dilution of a saturated solution which had been brought to pH7 by the addition of 5 ml. N NaOH per litre. The use of buffers was avoided because they are ineffective at very high salt concentrations.

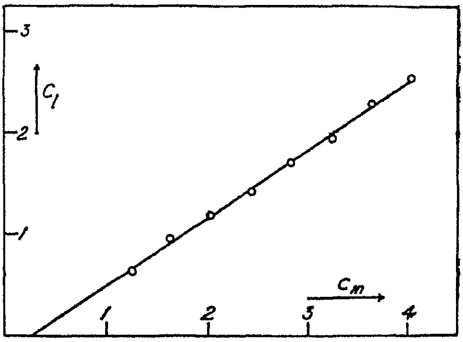


FIG. 2.—Ammonium sulphate concentration in the liquid of crystallisation (c_l) plotted against that in the suspension medium (c_m). No significance need be attached to the fact that the line does not pass through the origin, as the amount of deviation is within the experimental error.

No alteration in unit cell dimensions was observed over the entire range, although any lattice change exceeding 0·5 % could easily have been detected. The density of the wet crystals was found to be a linear function of salt concentration (Fig. 1). The unit cell volume of the normal wet crystals is 350,000 A.^{3}; by subtracting the volume of two hæmoglobin molecules it is found that 52·4 % of the cell volume are taken up by liquid of crystallisation. Fig. 2 shows the relationship between the salt concentration in this liquid, and that in the suspension medium, calculated from formula (1). Hydration of the protein is indicated by the fact that c_l is always lower than c_m. Fig. 3 shows the composition of the protein crystal as a whole, plotted against salt concentration in the suspension medium; it is seen that any salt which penetrates the crystal displaces an equivalent volume of water. In consequence the *total water content* of the crystals varies between 0·82 g. per g. protein in the absence of salt and 0·69 g. in saturated ammonium sulphate solution, while the *total volume of liquid* in the crystal remains the same. Fig. 4 shows the protein hydration as derived from equation (3). The hydration is seen to be of the order of 0·3 g. H_2O/g. protein throughout the whole concentration range studied. Salt concentrations near to the saturation point seem to have a slight dehydrating effect, otherwise the hydration appears to remain constant. (The unusually high value of w at c_m = 1·25 may be due to the presence of liquid-filled cracks in the crystals which showed signs of dissolving at this low concentration.)

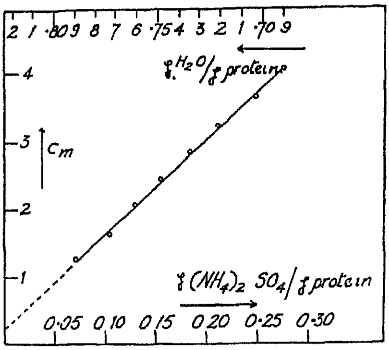


FIG. 3.—Composition of a horse methæmoglobin crystal as a function of c_m.

The Effect of Hydrogen Ion Concentration.

The effect of *p*H was studied at a constant salt concentration of 2·5 M to prevent dissolution of the crystals. Ammonium sulphate solution was again used as the basic medium whose *p*H was varied between 3·6 and 10·0 by replacing $(NH_4)_2SO_4$ with different quantities of mono- and diammonium phosphate, and, for *p*H values above 8, by adding concentrated ammonia. The *p*H determinations were made with the help of a glass electrode by kindness of Dr. E. F. Gale at the Department of Biochemistry.

The unit cell dimensions of the methæmoglobin crystals were not affected by immersion in media between *p*H 6 and 10. At *p*H 5·4 the lattice expanded by 7·3 %, and this expanded form in turn remained stable between *p*H 5·4 and 3·9. The expansion could be reversed by returning the crystals to a more alkaline medium. At *p*H 3·6 the crystals broke up into an amorphous coagulate. Their density changed little with *p*H; Fig. 5 shows protein hydration to be at a maximum near the isoelectric point and to diminish on either side.

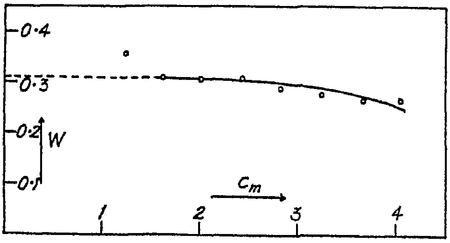


FIG. 4.—Hydration of protein in the crystal in g. H_2O/g. protein (w) as a function of c_m.

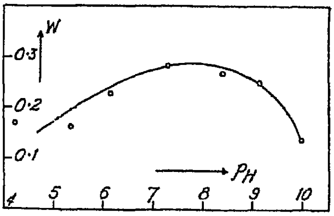


FIG. 5.—Hydration of protein in the crystal as a function of *p*H.

The unit cells of the normal and expanded crystals are seen in Fig. 6 (No. 5 and 7) as part of a comprehensive diagram of the unit cells of methæmoglobin crystals at all stages of swelling and shrinkage. Swelling at *p*H 5·4 is accompanied by a change of β from 111° to 84·5°; * the a

---

* The surprising change from the obtuse to the acute angle is genuine, and was confirmed by recording the X-ray pattern from a crystal as it changed from the normal to the expanded form. Comparison of the vector projections of the two forms also supports this observation.

and b axes remain constant in length and c hardly alters, while the layer spacing $c \sin \beta$ expands from 50·7 to 54·4 A. The perfection of the X-ray pattern is not affected by these lattice changes.

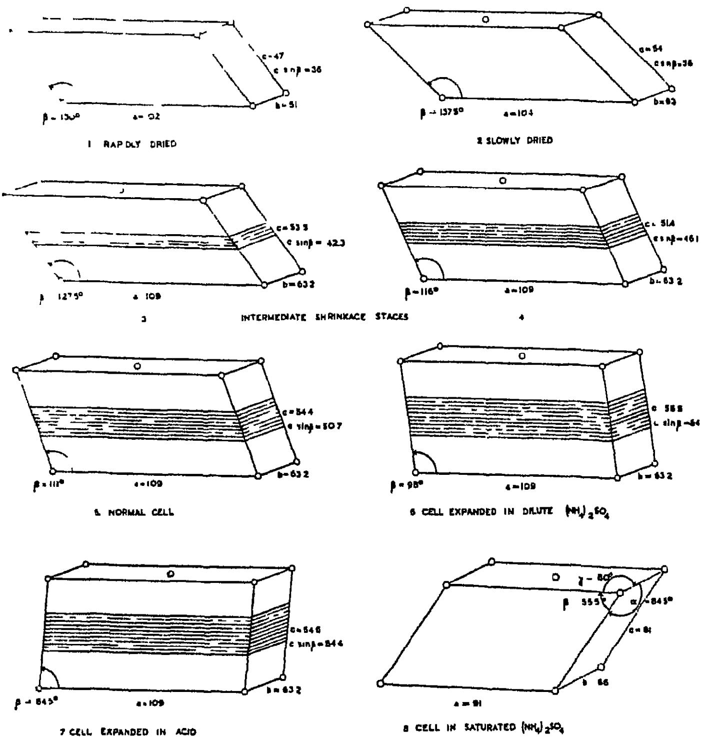


FIG. 6.—The diagrams show the unit cell dimensions of horse methæmoglobin crystals at all stages of swelling and shrinkage. The shaded parts in the centre of some of the parallelopiped indicate the thickness of the liquid layer. 1 mm. = 4 A.

The Effect of Pure Ammonium Sulphate Solution.

The *p*H of pure ammonium sulphate solutions decreases with increasing concentration. As salt concentration and *p*H vary simultaneously, the behaviour of methæmoglobin crystals immersed in a series of un-neutralised $(NH_4)_2SO_4$ solutions is more complex than in the experiments just described. The unit cell dimensions were normal only in the neighbourhood of a salt concentration of 2·8 M. Above this concentration the crystals changed into the acid-expanded form (Fig. 6, No. 7) already mentioned; this in turn became unstable in saturated $(NH_4)_2SO_4$ solution where the crystals assumed a triclinic form accompanied by a deterioration in the diffraction pattern and a marked increase in density (Fig. 6, No. 8). Below 2·8 M salt the crystals expanded by 8 % to a new and different unit cell, which remained stable up to 1·6 M. Dissolution of the crystals prevented observations at lower concentrations.

The triclinic cell is seen in Fig. 6, No. 8; its density is 1·28, indicating a hydration of only 0·04. Diagram No. 6 shows the unit cell which is stable between 2·8 and 1·6 M; it is similar to the acid-expanded one, except that β is 98° instead of 84·5°. This expansion, too, can be reversed by returning the crystals to a 2·8 M $(NH_4)_2SO_4$ solution; the whole process of swelling and reshrinking can be carried out without any adverse effect on the X-ray diffraction pattern.

Shrinkage.

Methæmoglobin crystals were made to shrink in stages (cf. [3]), by drying them extremely slowly. The two intermediary stages observed between the normal wet and the air-dried state are shown in Fig. 6, No. 3 and 4. The lengths of *a* and *b* remain constant while β changes first to 116° and later to 127·5°, involving a shrinkage of $c \sin \beta$ by 4·6 and 3·8 A. respectively. The cell dimensions of the air-dried crystals vary according to the method of drying, slowly dried ones having a larger unit cell and more perfect X-ray pattern than rapidly dried ones (Fig. 6, No. 1 and 2). Slow drying evidently preserves the alignment of the hæmoglobin molecules in a comparatively loose structure, while rapid drying produces closer packing, but destroys the phase relationship between corresponding groups of atoms in neighbouring molecules. The first shrinkage stage can be reversed by rewetting the crystals; no attempt at reversing the second stage was made, as it was only obtained once. The air-dried crystals tend to break up on rewetting, and a reversal of the shrinkage is therefore not feasible. The dried crystals can, of course, be dissolved and recrystallised.

Mechanism of Swelling and Shrinkage.

Neglecting such extreme treatments as rapid drying in air or immersion in saturated ammonium sulphate solution, a study of Fig. 6 shows that all swelling and shrinkage is practically confined to changes in the *c* spacing, brought about by a change in β from 84·5° in the acid-expanded crystals to 137·5° in the slowly dried ones. Of the total change of 53°, 32° at least can be produced reversibly. Detailed analysis of the X-ray pattern shows that the hæmoglobin molecules form coherent sheets which remain parallel to the *c* plane throughout the entire range of swelling and shrinkage. The sheets of hæmoglobin molecules are sheared parallel to each other as the angle β changes, but no alteration either in the thickness of the hæmoglobin molecules or in their internal structure seems to occur; it is only the thickness of the liquid layers between them that alters. Swelling and shrinkage in these crystals, therefore, is an inter- and not an intra-molecular process.

All swelling and shrinkage proceeds in discrete and reproducible steps. The lattice seems to click from one equilibrium position to another, while all intermediary positions are too unstable to be observed, at any rate on diffraction photographs requiring not less than one hour's exposure. In terms of the spacing between two neighbouring layers of hæmoglobin molecules (equivalent to $c \sin \beta$), the magnitude of the steps between the different equilibrium positions varies between 3·7 and 4·6 A., with the exception of the last step of drying which amounts to 6 A. These values were all obtained in ammonium sulphate solutions; other suspension media may well give rise to different equilibrium positions.

Discussion.

The lattice dimensions of wet methæmoglobin crystals are independent of the concentration of neutral electrolyte present in the suspension medium. They are, moreover, independent of pH over wide ranges, and only at one specific pH does a definite and reversible lattice change occur. It must be concluded, therefore, that the intermolecular forces which operate in protein crystals are of a totally different character from the long range forces present in gels of tobacco mosaic virus or bentonite. The lattice change which takes place around pH 5·5 provides at least a pointer to the type of force which may predominate. A survey of the dissociation constants of the amino-acid residues with polar side chains shows that only the imidazole group of histidine has a pK value in the region between 5 and 7; in pure histidine solutions its amino group carries a positive charge below *p*H 6·0 and is uncharged above it[5]; in peptides or proteins and in the presence of salts the *p*K value may be shifted considerably. Since there are 33 histidine residues in one hæmoglobin molecule, it may reasonably be suggested that the charging of their NH groups so alters the balance of electrostatic forces on the surface of the protein molecules that the lattice energy "troughs" undergo a substantial shift. The swelling could be explained as a result of the repulsive forces introduced by the appearance of the positive charges. Since no *p*K of any other ionised group is within the region investigated (3·9-10·0), it may be suggested that the absence of any lattice change other than that around *p*H 5·5 is due to the comparative constancy of the charge distribution on the hæmoglobin molecules in the *p*H ranges 3·9-5·4 and 5·6-10·0, say.

It is interesting to note that the transition from acid to alkaline methæmoglobin which takes place around *p*H 8·2 and is accompanied by a transformation of the absorption spectrum produces no lattice change. It was also found by allowing NaN_3 to diffuse into the crystals that the methæmoglobin could be transformed into azide-methæmoglobin, involving the attachment of an azide group 4 A. long to each of the four iron atoms, without change in cell dimensions. Since it is improbable that four such large groups could be tucked away in the interior of the protein molecule without causing structural changes affecting the X-ray pattern, it seems likely that the hæms lie on the surface, nevertheless the iron atoms, at any rate, can play no part in intermolecular binding, since otherwise the attachment of the azide groups would be expected to cause a change in cell dimensions.

The discontinuous character of all swelling and shrinkage may have two alternative explanations. Either the surface structure of the protein molecules is such that only certain distances between successive layers of protein molecules are structurally stable, or the liquid between those layers has itself a laminated structure, such that only one layer of salt solution can be added or removed at a time. The latter hypothesis is perhaps the more plausible one, though at the moment there is no evidence to distinguish between them. Hendricks *et al.*[6] found montmorillonite particles to incorporate several distinct layers of water of 3 A. thickness. The thickness of the layers in hæmoglobin crystals would have to be about 4 A., since this is the order of magnitude of the steps which separate successive states of swelling. This increased thickness compared to the pure water layers found in montmorillonite may be due to the presence of the sulphate ions; it remains to be seen whether ions of different size would alter the magnitude of the steps.

Since swelling and shrinkage is a purely intermolecular process, and does not affect the internal structure of the hæmoglobin molecules, hydration should be confined to the surface. X-ray data indicate that the hæmoglobin molecules correspond roughly to cylinders with an ellipitical base having a total surface area of 11,990 A^2. Supposing this area were covered with a layer of water one molecule thick and a density of 1·0, the weight of that water layer would be 31,000 g. per mole or 0·28 g. per g. protein, which agrees with the measured hydration of 0·3 g. per g. protein. There is no conclusive evidence as yet whether the additional intracrystalline liquid which has the same composition as the suspension medium is of a purely interstitial character. The fact that its composition can vary within such wide limits suggests that it is, and that its presence cannot easily be ascribed to protein hydration, as McMeekin and Warner have done in the case of salt-free lactoglobulin crystals.[2]

It should perhaps be pointed out that our results differ from McMeekin and Warner's in one important respect. They state that the total amount of water in β-lactoglobulin crystals is the same in the absence of salt and in the presence of high concentrations of ammonium sulphate. Since salt diffuses freely into the crystals, they have to assume that the space for the salt is provided by expansion of the lattice, an assumption which is not supported by any of their observations. Our experiments, on the other hand, prove that the lattice is rigid, and that any neutral electrolyte penetrating the crystal therefore displaces an equivalent volume of water; thus it is the total volume of liquid in the crystal and not the total water content which remains constant.

---

[5] Cohn and Edsall, *Proteins, Amino-Acids and Peptides* (New York, 1943).

[6] Hendricks, Nelson and Alexander, *J. Amer. Chem. Soc.*, 1940, 62, 1456; Hendricks and Teller, *J. Chem. Physics*, 1942, 10, 147.

For reasons of space, the details of the X-ray evidence are not discussed in this paper, and will be described in a later publication elsewhere. I have to thank Dr. Edna Davidson, Miss Joy Boyes-Watson and Miss Olga Weisz, who have helped me with certain of the experiments, Dr. H. Carlisle, who measured the unit cell of azide methæmoglobin and Dr. J. C. Boursnall, who kindly analysed some of the suspension media. I am also grateful to Dr. Gale for carrying out the *p*H determinations. The research was financed by a grant from the Rockefeller Foundation.

Conclusions and Summary.

The composition and swelling properties of horse methæmoglobin crystals were investigated by combining density measurements with X-ray diffraction studies. The following conclusions were reached.

(1) 52·4 % of the volume of normal wet methæmoglobin crystals consists of liquid whose composition can be varied within wide limits.

(2) It was shown by immersing crystals in a series of ammonium sulphate solutions at *p*H 7 that the unit cell dimensions and the general intensity distribution in the X-ray diffraction pattern are independent of the concentration of neutral electrolyte in the suspension medium. Similarly, they are independent of *p*H if the salt concentration is kept constant, except at *p*H 5·4, when the unit cell swells by a definite amount on acidification. In pure ammonium sulphate solutions any change in concentration involves a change in *p*H; more complex swelling effects were therefore observed.

(3) In methæmoglobin crystals layers of hæmoglobin molecules alternate with layers of liquid. Swelling and shrinkage produce variations in the layer spacing and shearing of the layers. All lattice changes proceed in definite, reproducible steps, involving changes in the layer spacing of the order of 4 A. at a time. The thickness and structure of the protein layers remain unaltered during swelling and shrinkage.

(4) At *p*H 7 the salt concentration in the liquid of crystallisation is a linear function of the salt concentration in the suspension medium; in ammonium sulphate solution the former is about two-thirds of the latter. If it is assumed that the liquid of crystallisation consists of two components, namely, "bound water" which is not available as solvent to mobile ions and "free liquid" through which ions can diffuse and which has the same composition as the suspension medium, it can be shown that protein hydration' ("bound water") amounts to 0·3 g. H_2O per g. protein. At the isolectric point hydration is largely independent of electrolyte concentration; it is, however, a function of *p*H and decreases on both sides of the iso-electric point. It is shown that this hydration is of the right order of magnitude for a monomolecular layer of water molecules covering the surface of the hæmoglobin molecule.

(*Molteno Institute and Cavendish Laboratory, Cambridge.*)

GENERAL DISCUSSION *

Dr. G. S. Hartley (*Maidenhead*) said: He had accumulated a considerable amount of evidence prior to 1940 which led to the conclusion that, in clear mobile solutions of paraffin chain salts, the paraffin arrangement in the micelles was essentially chaotic and liquid. This evidence had all been published. In a recent summary [1] the energy and probability aspects of this type of aggregation had also been considered. It was unnecessary therefore to repeat it, but one piece of evidence seemed in retrospect more significant than this very brief mention had suggested. This is best brought out in the case of the salt cetyl pyridinium chloride which comes out from aqueous solution in clusters of large needle crystals which are heavier than water. Nevertheless the solution is lighter than water. Thus an expansion of the total volume of the system, salt/water, occurs on solution; with all simple electrolytes a contraction occurs. The expansion can therefore only be produced by the behaviour of the paraffin part of the salt. This implies fusion of the paraffin crystal lattice, and indeed the expansion is quantitatively in approximate agreement with this view.

The need for a reconciliation between the liquid micelle and crystalline plate-like micelle theories might be avoided by mentioning that nearly all the author's measurements referred to concentrations below N/10. The X-ray evidence was obtained in higher concentrations and in fact results at the lower concentrations were negative.

It was, however, worth suggesting a more constructive reconciliation than this. In discussing the mechanism of preference for the spherical liquid micelle,[2] the author had concluded that the balance between this and an extended ellipsoidal micelle was rather delicate. In concluding that the liquid sphere would normally be preferred, the agreement was based on consideration of solutions so dilute that there could be no geometrical interference due to difficulty in packing of spheres. When, as in concentrated solutions, deformation due to packing requirements would have to be considered, the balance might well tend towards the formation of oblate spheroids. These could be quite adequate to explain the X-ray results and not give rise to the difficulty raised by Prof. Bernal about the mechanism of arbitrary interruption of the crystalline plates which Dr. Dervichian had to assume to accommodate all the water in the system. Incidentally this difficulty is a strong argument already raised by the author [3] in favour of the stability of the isolated spherical liquid aggregate.

Experimental evidence for the sensitivity of this balance is contained in the extremely specific behaviour of paraffin chain salts with regard to viscosity. In the case of hexadecoxy-benzene-4-sulphonate, substitution of copper for potassium transforms a solution mobile up to high concentrations into one which is strongly elastic even at so low a concentration as 0·02 %. No one can doubt that this mechanical behaviour requires thread-like aggregates which are probably of the myelenic form described by Dr. Dervichian.

Mr. V. R. Gray (*Manchester*) said: The concepts of gels and jellies here introduced are considerably modified by conclusions reached in parallel investigations on aluminium soaps made at Cambridge.[4] Studies of streaming birefringence, viscosity, and chemical composition of these substances have led to the conclusion that they are in fact high polymeric substances where the linkage between the monomer units is weaker than is commonly found for high polymers, and so the chain length can be altered by change of concentration or by small quantities of strongly co-ordinating substances (peptisers). The type of structure propose by us for, say, an aluminium "disoap" is as follows:

---

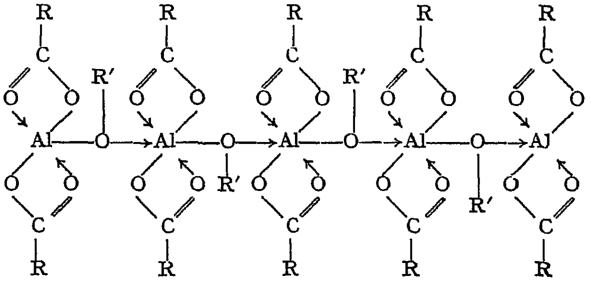
* On five preceding papers.

[1] Cp. Gregg, *J. Chem. Soc.*, 1942, 696.

[2] *Kolloid Z.*, 1939, 88, 22.

[3] *Proc. Roy. Soc. A.*, 1938, 168, 420.

[4] Forthcoming publications.



where the aluminium atoms are surrounded by octahedra of oxygen atoms, the octahedra being joined in chains in the manner observed in certain alumino-silicate minerals, and similar to the silicone resins.

In consequence of these conclusions, the "jelly" phase of the soap solution is exactly analogous with a concentrated solution of a high polymer, say a rubber-benzene solution, or nitrocellulose in acetone. Indeed the resemblance between these solutions is very striking, and is more than superficial. The observations on jelly-gel transitions for aluminium soap solutions thus have an exact counterpart in the swelling and solution of high polymers in solvents and swelling agents.

Prof. J. D. Bernal (*London*) said · The phenomena described particularly by Dr. Dervichian seem all to relate to the stability of smectic liquid crystals on the model first put forward in this room at an earlier Faraday Conference by Sir William Bragg.[5] Sir William pointed out that for a three-dimensional structure built out of equidistant layers, minimum energy structures must be closed. Plane structures would always have larger effective surfaces and hence higher energy. He showed that these structures all belong to the class of Dupin's cyclides, of which one particular case is the set of infinite coaxial cylinders, the myelinic structures of Dervichian. In the light of this I feel that a soap micelle in the form of relatively thick plates, as suggested by Dr. Dervichian, would probably be unstable. It seems likely also that the micelles are not spherical, both because the spherical form is not particularly stable owing to the impossibility of accurate packing at the centre, and because sols which are spherical do not show the phenomena of characteristic soap sols. The X-ray evidence so far on soaps only gives, and that rather roughly, the mean distances between layers. The actual arrangements of these layers have not been determined, but might be so by more careful optical and X-ray studies. At any rate they deserve to be studied as much as any earlier known forms of liquid crystals. These remarks, of course, do not apply to the relatively coarse and ribbon-like crystals that appear in low temperature gels in soap systems.

I would like to add that batonnets and tactoids should not be confused. They have different degrees of molecular symmetry. In batonnets, which occur frequently in long chain molecule systems, there is ample evidence of parallel layer formation. Tactoids appear to be of two kinds, one of which has been most carefully studied, where long particles of unequal length lie in parallel orientation, as in nematic liquid crystals, but where there is in addition a two-dimensional regularity at right angles to the particle lengths. A second kind consists of platy particles with similar orientation but where the equidistance is probably only one dimension, giving rise to a structure of orthorhombic symmetry.

---

[5] *Trans. Faraday Soc.*, 1933, 29, 1056.

RANDOMNESS IN THE STRUCTURES OF KAOLINITIC CLAY MINERALS.

BY G. W. BRINDLEY AND K. ROBINSON.

Received 21st October, 1946.

1. The Kaolin Minerals.

Four distinct minerals, nacrite, dickite, kaolinite and halloysite, have been recognised having structures formed by the superposition of layers with the ideal composition $Al_2Si_2O_5(OH)_4$ or, in terms of oxides,

$$Al_2O_3 . 2SiO_2 . 2H_2O.$$

There exists also a hydrated form of halloysite, sometimes termed endellite,[1] with the ideal composition $Al_2O_3 . 2SiO_2 . 4H_2O$, which readily changes to the dehydrated form with $2H_2O$ per "molecule"; the process is largely irreversible unless special methods are used such as MacEwan[15] has described. The number of structural layers per unit cell decreases from six in nacrite (Hendricks[8]) to two in dickite (Gruner,[7] Hendricks[9]), and one in kaolinite (Brindley and Robinson[2]). In dehydrated halloysite the layers are stacked more or less regularly along the *c*-axis, but with largely random orientations in their own planes (Alexander *et al.*[1]; Brindley, Robinson and MacEwan[4]). In hydrated halloysite the kaolin-type layers are interspersed with single sheets of water molecules (Hendricks[9, 11]).

The structure of kaolinite was first analysed by Gruner,[6] who considered it to be monoclinic with two structural layers per cell but the present writers[2] have subsequently shown it to be triclinic, with the cell dimensions $a_0 = 5{\cdot}14$ A., $b_0 = 8{\cdot}93$ A., $c_0 = 7{\cdot}37$ A., $\alpha = 91{\cdot}8°$, $\beta = 104{\cdot}5°$, $\gamma = 90°$. This cell is approximately monoclinic, and contains one layer per cell. While the new structure does not radically change our conception of kaolinite itself, it suggests a new point of view regarding the relationship of kaolinite to dehydrated halloysite; the two minerals have almost the same basal spacing of 7·1-7·2 A. corresponding to the thickness of one layer, but whereas all layers are similarly orientated in an ideal kaolinite, they have completely random orientations in an ideal halloysite. The possibility may therefore be considered that between these extremes there may exist minerals with various degrees of randomness in the ordering of the layers. Evidence has recently been obtained by the writers[3] that certain fireclays may occupy an intermediate position, while evidence is now accumulating that various degrees of order may exist among minerals normally classed as kaolinites.

---

[1] Alexander *et al.*, *Amer. Min.*, 1943, **28**, 1.
[2] Brindley and Robinson, *Nature*, 1945, **156**, 661; *Min. Mag.*, 1946, **27**, 242.
[3] Brindley and Robinson, *Trans. Brit. Ceramic Soc.* (in press).
[4] Brindley, Robinson and MacEwan, *Nature*, 1946, **157**, 225.
[5] Grimshaw, Heaton and Roberts, *Trans. Brit. Ceramic Soc.*, 1945, **44**, 69.
[6] Gruner, *Z. Krist.*, 1932, **83**, 75. [7] Gruner, *ibid.*, 394.
[8] Hendricks, *ibid.*, 1938, **100**, 509.
[9] Hendricks, *Amer. Min.*, 1938, **23**, 295.
[10] Hendricks, *Physical Rev.*, 1940, **57**, 448.
[11] Hendricks and Jefferson, *Amer. Min.*, 1938, **23**, 863.

The development of random structures in kaolin minerals is connected with the structure of the layer itself shown in projection in Fig. 1; it consists of a sheet of Si—O tetrahedra, Fig. 1(*a*), on which rests a sheet of Al—O, (OH) octahedra, Fig. 1(*b*), and the complete layer as it appears in kaolinite is shown in Fig. 1(*c*), the heights of the atoms above the basal plane being indicated in the key. The Al sheet can rest on the Si sheet in six different ways indicated diagrammatically in Fig. 2 and, with the *a*- and *b*- axes as shown, Fig. 2(*a*) corresponds to kaolinite. The six arrangements fall into two groups, (*a*), (*b*), (*c*) and (*d*), (*e*), (*f*), the members of each group differing only in the location of the aluminium atoms. When layers of each kind are placed above each other, displaced in accordance with the direction of the *c*-axis of kaolinite, i.e., with $\alpha = 91{\cdot}8^\circ$, $\beta = 104{\cdot}5^\circ$, then the basal oxygens of one layer are related to the uppermost hydroxyls of the layer beneath in the manner shown in Fig. 3 where (i) corresponds to the stacking of layers 2(*a*), (*b*) and (*c*), and (ii) corresponds to 2(*d*), (*e*) and (*f*). In kaolinite (Fig. 2(*a*) and 3(i)), as in nacrite and dickite (Hendricks [8]), the interlayer binding arises from a hydrogen bond between closely associated pairs of O and (OH) atoms, and this, in fact, gives a simple explanation for the value found for the β-angle. It is clear from Fig. 3 that only layers of types 2(*a*), (*b*) and (*c*) will give this kind of bonding between pairs of O and (OH) atoms when the angle β is 104·5°.

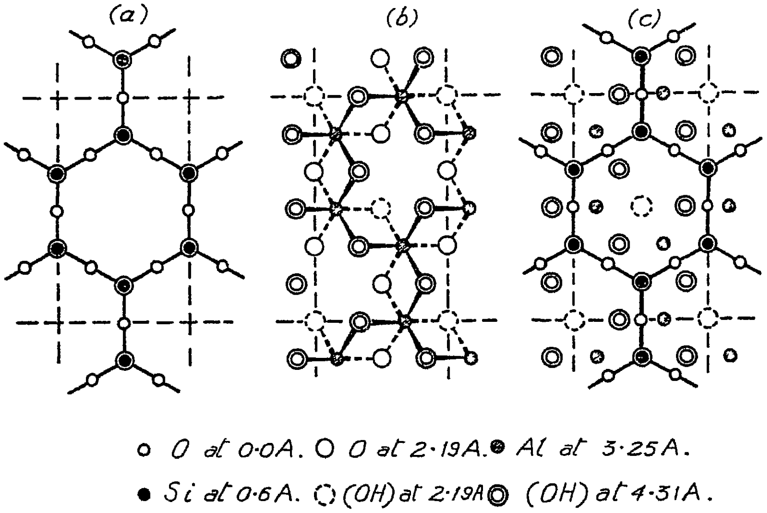

(a)
(b)
(c)
o O at 0·0A. O O at 2·19A. ● Al at 3·25A.
• Si at 0·6A. ⦾ (OH) at 2·19A ◎ (OH) at 4·31A.

FIG. 1.—Structure of the kaolin layer projected on (001). (*a*) Si-O sheet; (*b*) Al—O (OH), sheet; (*c*) Complete layer Heights of atoms in A., above basal (001) plane.

In addition to random features within the kaolin layers, the layers themselves may be displaced randomly with respect to each other by integral multiples of $b_0/3$ in the *b*-direction as Hendricks [10, 12] in particular has pointed out, such displacements being possible because the uppermost (OH) groups are repeated at intervals of $b_0/3$. That displacements of a still more random character are possible is shown by the appearance of many of the reflections from halloysite which can be interpreted in terms of completely or almost completely random orientations of the layers in their own planes, i.e., with respect to both *a*- and *b*- axes. We have recently found evidence in certain fireclays for random displacements parallel to the *b*-axis only. Evidence is also forthcoming for some variation

---

[12] Hendricks and Jefferson, *Amer. Min.*, 1939, 24, 729.

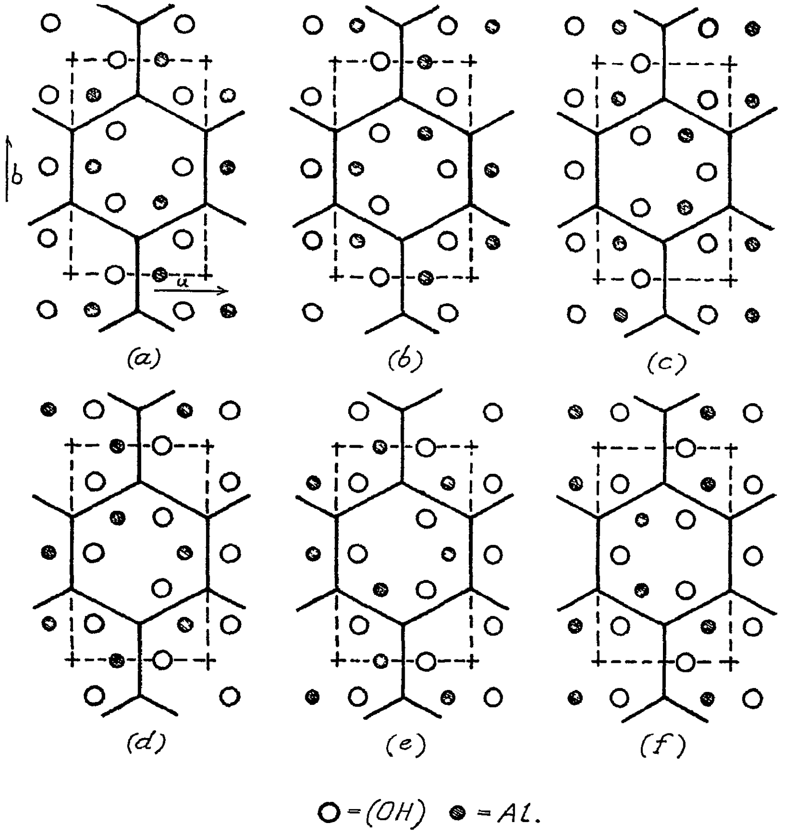


FIG 2.—Diagrams showing the six ways in which the Al—O, (OH) layer can be superposed on the Si—O layer.

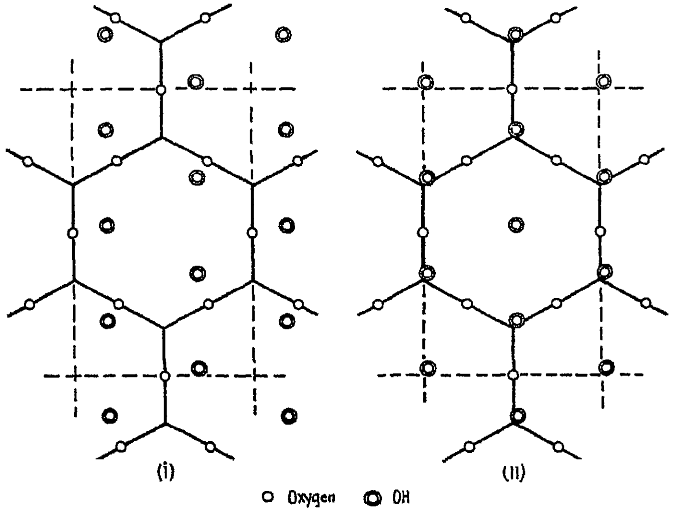


FIG. 3.—Diagrams showing the relations between the basal O atoms and uppermost (OH) groups of two adjacent kaolin layers; (i) Corresponds to Fig. 2(*a*), (*b*), (*c*), and (ii) to Fig. 2(*d*), (*e*), (*f*).

in the stacking of the layers above each other, producing a broadening of the $(00l)$ reflections, but the variation of lattice spacing in dehydrated halloysite is not on so large a scale as may arise in montmorillonite where one or more water layers can be interposed irregularly between the silicate layers (see Hendricks and Teller[13]).

2. X-Ray Powder Data for Kaolin Minerals.

The general characteristics of the X-ray powder photographs we have obtained are shown diagrammatically in Fig. 4 as far as the (060) line; the lower order lines are practically the only ones corresponding to single reflections even when a 20 cm. diameter camera is employed. A progressive variation is seen on passing from a good kaolinite, (*a*), to a typical dehydrated halloysite, (*d*); (*b*) corresponds to a variety of kaolinite showing some features found in (*c*) and (*d*), while (*c*) corresponds to a number of kaolin-type fireclays from the Coal Measure series in Yorkshire. The main features to be considered are the growth of the band-like effect from (*a*) to (*d*) and the changes in the number and character of the lines.

(i) Randomness in the Structure of Dehydrated Halloysite.—The band-like effect is the major feature of Fig. 4(*d*), and appears also on the (020) line of (*b*) and (*c*). In halloysite its origin is the high degree of randomness in the orientation of successive layers. The observed reflections can be indexed as $(00l)$ and (hk), the latter having the diffuse character of two-dimensional X-ray reflections. The theory of reflection by two-dimensional gratings has been treated by von Laue[14] and by Warren.[16] Fig. 5 shows preliminary results relating to the shape of the (02) reflection; the full lines are photometric curves of a Missouri dehydrated halloysite before and after extended heat-treatment—discussed in § 3—while the dashed line is calculated from Warren's expression. The close similarity in shape of the observed and calculated curves is further evidence of the randomness of the layers.

(ii) Randomness in a Kaolinitic Fireclay Mineral.—The structure of the clay mineral in a number of fireclays giving X-ray photographs of type 4(*c*) has recently been considered by the present writers[3] in some detail, and it will suffice to set down here the main conclusions. In the first place the experimental evidence when carefully considered is against a diagram of type 4(*c*) arising from a kaolinite-halloysite mixture. Such a mixture would have to contain about 90 % halloysite, and would show lines and bands not found with the fireclay; furthermore lines occur which cannot be attributed to such a mixture. The reflections can be indexed on the basis of a monoclinic cell with $a_0 = 5{\cdot}14$ A., $b_0 = 8{\cdot}91$ A., $c_0 = 7{\cdot}37$ A., $\beta = 104{\cdot}5°$; the indices are given in Fig. 4(*c*). The cell dimensions are therefore the same as for kaolinite except that $\alpha = 90°$ instead of $91{\cdot}8°$. The following explanation has been suggested for the apparent monoclinic symmetry. In the first place, although the *shape* of the unit cell in kaolinite is almost monoclinic, i.e. $\alpha \simeq 90°$, the symmetry of the atomic distribution is not monoclinic owing to the positions occupied by the Al atoms. It seems likely that $\alpha \neq 90°$, on account of the distribution of the Al atoms. If in the fireclay mineral arrangements (*a*), (*b*) and (*c*) of Fig. 2 are statistically distributed (and these arrangements only differ as regards the Al positions), then the small displacements of successive layers which produce the α-angle of kaolinite will now take place equally in three directions at 120° to each other, and therefore give zero resultant. The structure will then be pseudo-monoclinic. Direct proof of this kind of randomness in the positions occupied by the Al atoms is difficult to obtain,

---

[13] Hendricks and Teller, *J. Chem. Physics*, 1942, **10**, 147.
[14] von Laue, *Z. Krist.*, 1932, **82**, 127.
[15] MacEwan, *Nature*, 1946, **157**, 159.
[16] Warren, *Physical Rev.*, 1941, **59**, 693.

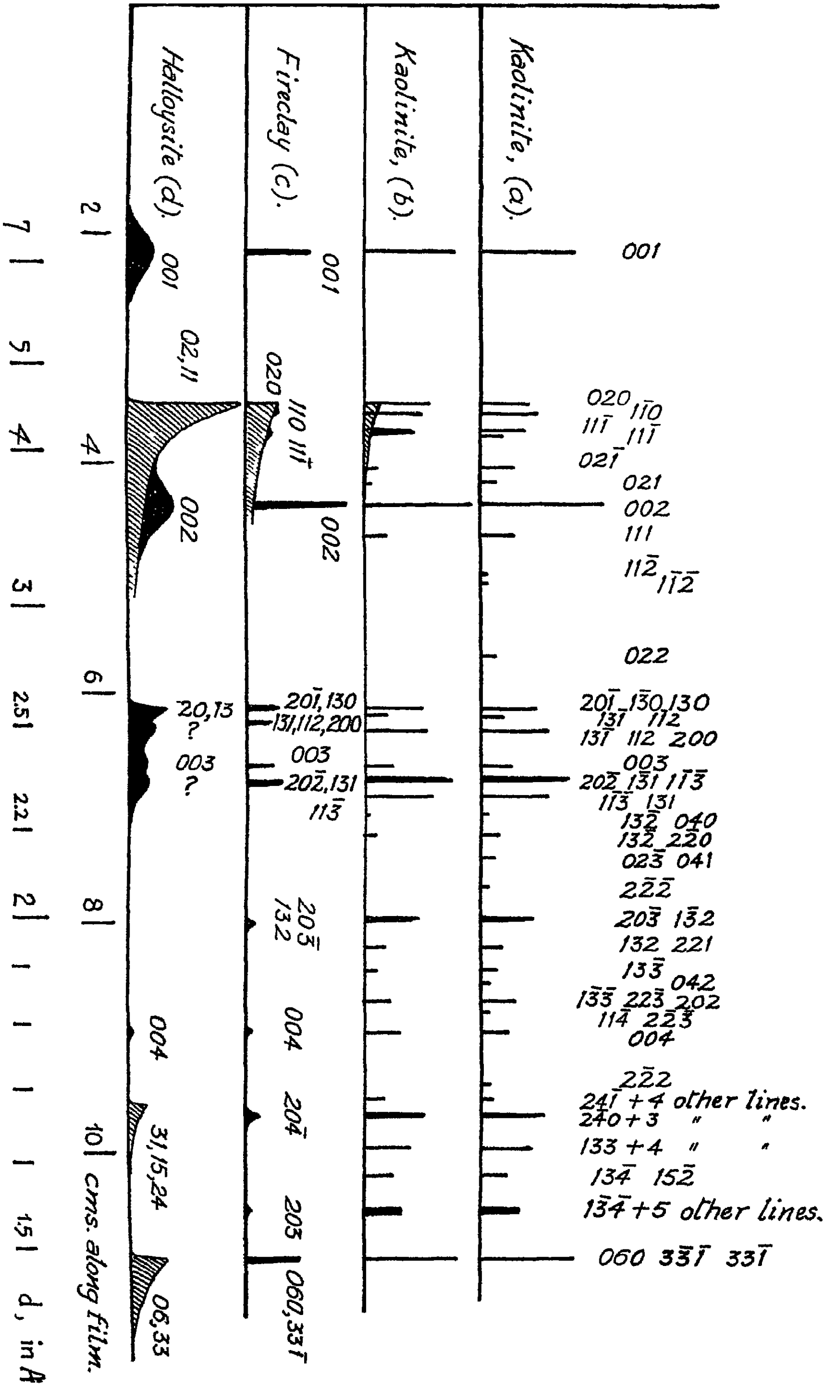


FIG. 4.—Diagrammatic representation of X-ray photographs of some kaolin minerals. The distances along the film, in cm., correspond to a camera of 20 cm. diameter.

since other kinds of randomness also occur which make any precise comparison of observed and calculated intensities practically impossible. We have examined this question for the lower order reflections and consider that the available evidence does not conflict with this hypothesis.

The indices of the reflections given in Fig. 4(*c*) are for the most part of a type with the *k* index an integral multiple of 3, which indicates displacements of layers by $nb_0/3$ parallel to *b*, but the existence of the weak (110) and $(11\bar{1})$ lines shows that these displacements are not of such frequent occurrence that they can be regarded as of strictly random value.

The band associated with the (020) reflection is an important feature of the diagram in Fig. 4(*c*), for it is the only certain band visible in this diagram. The evidence is against it arising from an admixture of halloysite. To account for the occurrence of this one band, it is necessary to suppose that random displacements of layers occur in the *b*-direction only and not, as in halloysite, in both the *a*- and *b*-directions. Randomness of this more restricted kind leaves reflections of type (*hOl*) unaffected, and in our previous paper[3] we have shown that it may well explain why the only observable band is associated with the (020) reflection. Consider, for example, the combined reflection $(20\bar{1})$, (130) which shows a band with halloysite but not with the fireclay mineral; $(20\bar{1})$ is a crystalline reflection unaffected by displacements parallel to *b* and (130) is also a crystalline reflection from those parts of the structure where displacements of $nb_0/3$ occur; random displacements parallel to *b* give the two-dimensional, banded reflection (13). Under such conditions the band will probably escape detection. The (02) band is likely to be prominent because the crystalline (020) reflection will be weak or non-existent ($k \neq 3n$), and the (02) band is known to be strong because of its pre-eminence in the halloysite diagram.

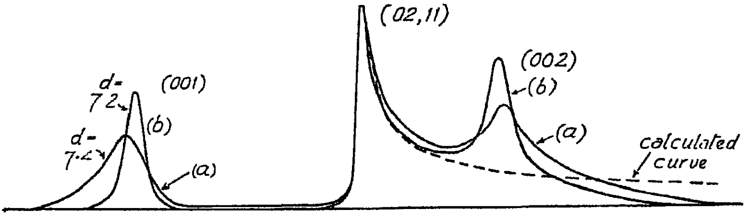


FIG. 5.—Microphotometer traces of X-ray photographs of (*a*) natural, (*b*) heat-treated Missouri Halloysite. The dashed curve is calculated from Warren's expression for the intensity diffracted by a two-dimensional crystal lattice.

We consider, therefore, that fireclays giving an X-ray diagram of type (*c*) in Fig. 4 contain various random features in their structures. Randomness in the Al positions produces a pseudo-monoclinic structure, while displacements of $nb_0/3$ parallel to the *b*-axis give crystalline reflections with the *k* index mainly a multiple of 3 and random displacements parallel to *b* give the characteristic band effect.

(iii) **Randomness in Kaolinite.**—Photographs of type (*b*) in Fig. 4 are given by many kaolinites; the main characteristics are the weak band associated with the (020) line and the lack of resolution of the $(11\bar{1})$, $(1\bar{1}\bar{1})$ doublet. In this case we cannot rule out the possibility that the material is a mixture of kaolinite and halloysite. The mixture, however, would have to contain about 30 % halloysite and the kaolinite would have to be of very small crystal size to explain the lack of resolution of the doublet. We have not so far found a band associated with a resolved doublet and conversely when the band is absent, the doublet is generally very clearly resolved. While these statements do not disprove the mixture hypothesis, they do suggest a close association of the band with the lack of resolution. Random shifts along the *b*-axis may produce the band and distortion of (*hkl*) planes arising from some randomness in the Al positions may blur the doublet. We cannot at present draw any more definite conclusion than that kaolinites of type (*b*) may represent an intermediate stage between completely ordered kaolinite and completely disordered halloysite, a stage which is much nearer the former than the latter.

3. The Dehydration Process of Hydrated Halloysite.

Chemical analyses of dehydrated halloysite, after thorough drying at 110° c. to remove adsorbed water, always show rather more water than the 13·9 % which corresponds to the ideal formula $Al_2O_3 . 2SiO_2 . 2H_2O$, while kaolinite gives results in good agreement with this value ; data given by Alexander *et al.*[1] indicate about 15 % water in dehydrated halloysite. Thermal analyses by Roberts and his collaborators[5] have shown an endothermic peak at 150-200° c. for halloysite which is not found with kaolinite ; this peak may arise from the additional water being held rather loosely in the lattice as compared with the 13·9 % which is accounted for by the composition of the kaolin layers.

X-ray diagrams show the basal (001) reflection of dehydrated halloysite to be about 7·4 A. and very broad (see Fig. 5). Heat treatment at temperatures from 100-350° c. sharpens this reflection and reduces the spacing to about 7·2 A., but never to a value quite so low as 7·14 A., the normal value for kaolinite.[4] A simple calculation shows that the increased spacing can be correlated with the excess water. Assuming 15 % water in the 7·4 A. halloysite, we have roughly 1 in 15 parts (i.e. about 7 %) more water than in kaolinite ; this additional water, if interposed between kaolin layers, would require one water layer to 15 kaolin layers. Assuming 7·14 A. and 3·0 A. respectively for the thickness of kaolin and water layers, we obtain the mean spacing of dehydrated halloysite to be

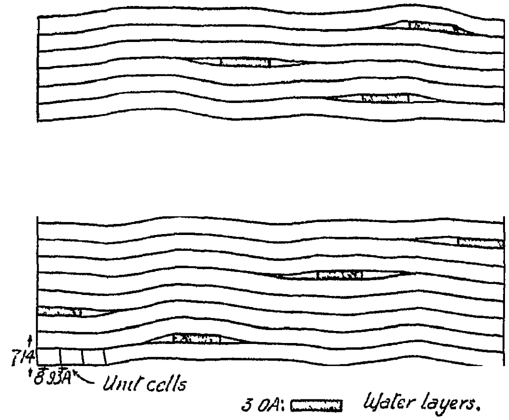


FIG. 6.—Possible arrangement of water inclusions in a partially dehydrated halloysite.

$$(15 \times 7{\cdot}14 + 3{\cdot}0)/15 = 7{\cdot}34 \text{ A.}$$

which is of the right order. We think that when fully hydrated halloysite is dehydrated, either naturally or very gently in the laboratory, e.g., over P_2O_5, some water may be trapped in the lattice in the manner illustrated diagrammatically in Fig. 6 ; as regards the size of the unit cells and the thickness of the layers, the diagram is drawn correctly to scale, but we have no evidence so far regarding the lateral extent of the water inclusions. It may be mentioned that the width of this diagram corresponds approximately to the value of *L* calculated by applying Warren's theory to the (02, 11) band in Fig. 5, namely 165 A. The distortion of the layers agrees generally with the broadness of the (*00l*) reflections observed before heat treatment. The (02, 11) band is not appreciably affected by heat treatment and therefore no growth of the two-dimensional gratings takes place.

MacEwan[15] has shown that organic molecules such as glycerol can enter halloysite after *very gentle* dehydration, but not after heat treatment. Presumably the former treatment leaves a fairly open structure which permits entry of the organic molecules, but heat treatment, by eliminating much of the excess water, " seals off " the layers which cannot then be penetrated.

4. Summary.

Randomness in the arrangement of kaolin layers in minerals of the kaolinite-halloysite type is discussed. It is suggested that kaolinite and halloysite (dehydrated form) may represent end-members of a series in which the disorder of the layers increases from kaolinite to halloysite. The ideal kaolinite has a perfectly ordered sequence of layers, and is triclinic in symmetry. The ideal halloysite has a sequence of layers completely disordered as regards their orientations. Fireclays are found which are pseudo-monoclinic probably owing to randomness in the positions occupied by Al atoms, and which have a considerable degree of randomness in the displacement of layers parallel to the *b*-axis. Kaolinites are found which show some of the features of these fireclays.

The process of dehydration of hydrated halloysite is discussed in relation to composition and the effect of heat treatment. It is suggested that dehydration is not usually complete, and that inclusions of water molecules trapped between the kaolin layers explain (i) the greater basal spacing and broadened basal reflections, (ii) the additional endothermic peak in thermal analysis curves, and (iii) the formation in some cases of organic-halloysite complexes.

Finally, we have pleasure in acknowledging a grant from the Department of Scientific and Industrial Research to provide an assistant (K.R.) for this work, and in recording our indebtedness to Dr. A. L. Roberts and his research team in the Fuel Department of this University for frequent discussions on the broader aspects of clay mineralogy, and their generous assistance in making thermal analyses of many of the materials we have employed.

Physics Laboratories,
University of Leeds.

ON THE HYDRATION OF MONTMORILLONITE.

BY J. MERING.

Received 27*th October*, 1946 *in French.*

Translated by S. M. *and* E. N. TIRATSOO.

The following facts have been established by a number of investigations, especially those of Hoffmann and Hendricks.

(1) The clay minerals of the montmorillonite group are built up of layer-lattices of the same structural model as those of which pyrophyllite crystals are composed, $Al_2Si_4O_{10}(OH)_2$.

(2) In montmorillonite these layers are not stacked regularly enough to form true crystals. They are grouped in small numbers—not exceeding a few dozen—forming stacks, inside which the only degree of organisation revealed by X-ray diffraction is a parallelism sufficient to result in selective (00l) reflexions. Such a stack of parallel layers will be called a *primary particle*. It will be seen that the " elementary unit-cell " of montmorillonite can only be defined by the two ortho-hexagonal translations $a = 5{\cdot}20$ A., $b = 8.96$ A., which characterise the periodicity in the plane of a single layer. This unit-cell contains 20 atoms of oxygen and four OH-groups.

(3) Contrary to what occurs in pyrophyllite, whose chemical composition is well defined, varied substitutions take place in the montmorillonite layer; in particular, Mg atoms replace some of the Al atoms in 8-co-ordinated positions, and Al atoms partially replace Si atoms in 4-co-ordinated positions.[1] The formula for montmorillonite can therefore be written approximately as

$$[Al_{(2-x)}Mg_xSi_{(4-y)}Al_yO_{10}(OH)_2]Na_{(x+y)},$$

$(x + y)$ is approximately 0·33, and represents the negative charge produced by the substitutions and compensated by the cations located outside the layers, chiefly in the spaces between them.

From the point of view of the substitutions alone, montmorillonite is thus situated between pryophyllite, whose build-up is neutral and not substituted, and the micas, in which the substitutions are more important and regular, bringing about a greater regularity in the distribution of the excess negative charge of the layers. The very irregular character of the distribution of the charges in montmorillonite weakens the attraction between the layers enough for water molecules to force them apart and enter into the spaces, making the whole surface of the layers accessible to hydration, and the compensating cations accessible to exchange.

(4) This phenomenon of hydration in montmorillonite is no doubt favoured by the small dimensions of the layers ; the work of forcing the layers apart is done by the water molecules present on their periphery, and the smaller the layer, the less work to be done by each water molecule.

(5) The water, which becomes fixed between the silicate sheets forms itself into layers.[2] In each layer the water molecules form a hexagonal pattern (like carbon atoms in graphite) conforming to the structure of the silicate. The number of superimposed layers [3] may be 1, 2, 3 or 4, resulting in basic distances of 12·4 A., 15·5 A., 18·4 A. and 21·4 A. between the sheets. The average number of layers varies continually and reversibly with the vapour pressure in the surrounding space. The curve of equilibrium depends on the nature of the exchangeable cations. Each primary particle has generally two different spacings corresponding to the simultaneous presence of two successive stages of hydration. In certain cases the two " hydrates " tend to separate within a single particle and the X-ray diagrams give the impression of two separate " phases." [3] But in most cases the distribution of the two different spacings occurs at random, and the apparent spacings observed on the X-ray photographs represent a statistical effect.

(6) Montmorillonite, when exposed to R.H. $> 50\ \%$, " fixes " quantities of water which are considerably greater than those which could be contained in the space between sheets indicated by X-rays. Therefore a part of the total reversible sorption takes place on the exterior of the primary particle, or at least outside the organised regions of the material. Let us call this exterior water *adsorption* water, and the water fixed between the elementary sheets *hydration* water. The term "constitution water" is reserved for the water produced by the elimination of OH-groups inherent in the elementary sheets.

In the study of the mechanism of water fixation, the factors inhibiting this property must be of particular interest. These can be classed in two groups according to whether they oppose the creation of the necessary sorption space or decrease the affinity of the silicate sheets for water. An example of the first is the fixing of the sheets by the " tie-across " effect due to diamines ; [4] an example of the second effect is the treatment with di-azomethane,[5] which renders the surface of the silicates hydrophobic by fixing CH_3-groups. The capacity for hydration of montmorillonite also disappears gradually as it is progressively heated. The dehydration of the mineral ceases to be reversible above 300° c., while at about 850° c. a totally inactive substance is obtained. This dehydration leads to the gradual elimination of OH-groups from the silicate sheets.

---

[1] Marshall, *Z. Krist. A*, 1935, **91**, 433-449 ; Nagelschmidt, *Min. Mag.*, 1938, XV, **162**, 140-55.

[2] Hendricks and Jefferson, *Amer. Min.*, 1938, **23**, 863-75.

[3] Bradley, Grim and Clark, *Z. Krist.*, 1937, **97**, 216-22.

[4] Hendricks, *J. Physic. Chem.*, 1941, **45**, 65-81.

[5] Berger, *Chem. Weekblad*, 1941, **38**, 42-3.

In this irreversible dehydration, we have sought the relationship between the absorption capacity and the number of residual OH-groups, taking into account the nature of exchangeable cations.

Experimental Method.

A few grams of montmorillonite are heated in an electric furnace at constant temperature to constant weight. From a fraction, the total "residual water" *R* remaining at the end is determined. The rest is rehydrated by leaving for 10 hours in distilled water, then dried, and if necessary pulverised. By varying the temperature of the treatment, a series of samples characterised by decreasing values of *R* can be prepared in this way. For each an isotherm of water sorption is determined by leaving it until it has attained equilibrium at a constant temperature of 20° c. in an evacuated vessel, in the presence of salt solutions giving well-defined R.H.'s.

For all the samples, one can follow with the help of X-ray diagrams, (*a*) the alteration of the mode of hydration, by examining them at various relative humidities; (*b*) the alteration in the structure of the silicate, by examining them in a *dry state*, i.e. when they have lost their water of hydration by being heated to 400° c. and are sealed in small Lindemann glass tubes. These observations have been made both for powdered and oriented specimens of Ca-montmorillonite and Na-montmorillonite, prepared from the pure mineral from Taourirt (Morocco), with Camp-Berteau montmorillonite and an American montmorillonite of the Black Hill type.

Decrease of the Hydratable Surface.

For Na-montmorillonite isotherms of all samples prepared from the same parent substance are of the same form: by adapting the scales of the ordinates they can be made to coincide. For Ca-montmorillonite this is only true of samples whose sorption capacity has been considerably reduced. It can therefore be stated that the irreversible dehydration of the mineral produces a progressive decrease of the hydratable *surface*, without noticeably affecting the law of equilibrium in the portion remaining active. The nature of this phenomenon is shown by the X-ray diagrams and the following points may be mentioned.

(*a*) The fundamental reflections (reflections from the plane of the layers) of the samples with a diminished hydration capacity never form —except in the dry state—a rational series (i.e. with values of sin θ proportional to integral numbers).

(*b*) If photographs taken at the same R.H. are compared, the intensity of part of the fundamental reflections diminishes and then vanishes as one gets further away from either the natural state or the totally inactive state obtained at about 850° c. The reflections from the original mineral which are preserved without much diminution of intensity are those whose apparent spacing is in the neighbourhood of an integer submultiple of 9·6 A. Thus the basic layers contain two kinds of elementary spacing: the fixed spacing of 9·6 A., which is that of the non-hydrated mineral, and variable spacings which are the same as in the original montmorillonite. In other words, the decrease of the hydratable surface is discontinuous and takes place layer by layer. As the treatment temperature is raised, an increasing number of layers become immobilised by cross-linking the rest of the layers being still able to move apart freely.

The Relationship between the Decrease of the Hydratable Surface and the Number of OH-Groups Eliminated.

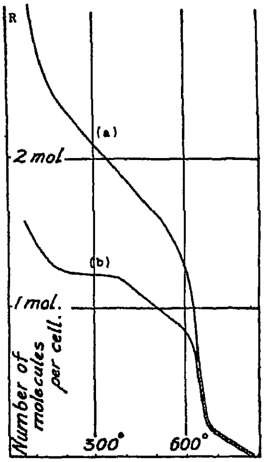


FIG. 1.

Some difficulty is encountered in determining the number of OH-groups present, because the dehydration curve of montmorillonite with varying temperatures is continuous (curve *a* of Fig. 1) and does not allow for any distinction between the water fixed by adsorption and hydration and the constitution water derived from OH-groups. We have nevertheless measured the quantity of constitution water in every case, using three different methods. All three methods are partly arbitrary, but their use is justified in that they give concordant results.

(*a*) The rehydrated sample is heated to different temperatures; after each heating the quantity of water R' remaining is determined, as well as the total sorption under constant conditions. If this total sorption is plotted as a function of R' the curve is horizontal at first and then falls abruptly. It is assumed that the sorption remains constant as long as the constitution water is not affected. The measurement of the latter is therefore given by R'_0 corresponding to the end of the horizontal portion of the graph. All determinations of the weight of water are made by weighing the samples after each operation. Applied to the original Ca-montmorillonite, this process fixes the number of OH groups at 4 per unit-cell in accordance with the structural formula.

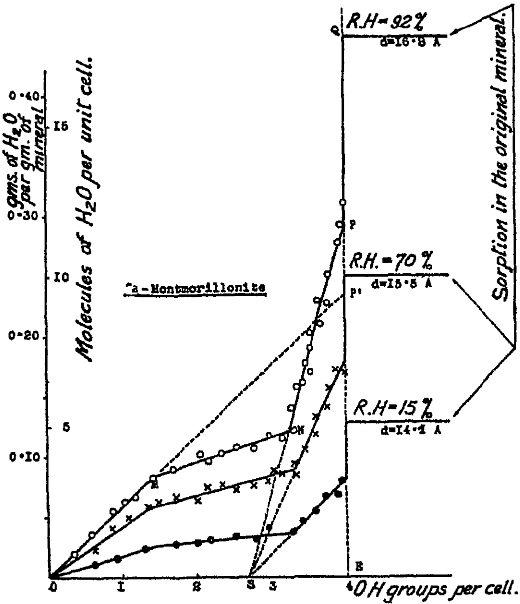


FIG. 2.

(*b*) The rehydrated sample is heated to increasing temperatures together with a sample of the original clay for comparison. When the total quantity of water has been reduced to a figure corresponding to four OH-groups in the check sample, the water now remaining in the test sample is assumed to represent its constitution water.

(*c*) The dehydration curves of samples whose hydratable surface is greatly reduced present a well-defined horizontal portion (curve *b* of Fig. 1). The ordinate of this can be taken as a measure of the quantity of constitution water. The agreement between the results obtained by the three methods is good for Ca-montmorillonite.

For the Na-montmorillonite series, only the results of methods (*b*) and (*c*) agree, method (*a*) producing much lower values. This shows that an Na-montmorillonite can be deprived of part of its OH-groups without diminishing its sorption.

For both series the OH-contents thus determined are above those corresponding to the values of residual water *R* measured after the first treatment in the furnace; i.e. the OH-groups which are eliminated by heating, are in part reconstituted when the samples are rehydrated. The proportion of reconstitution is of the order of 60 % at the beginning of each series, diminishes as the temperatures of treatment increase and becomes zero at about 850° c. Fig. 2 and 2*a* give the decrease of absorption capacity for Ca-montmorillonite and Taourirt Na-montmorillonite. The abscissæ represent the number of OH-groups per unit-cell and the ordinates the amounts of water reversibly fixed in different relative humidities. These quantities are expressed, as in the rest of this paper, in grams of water per gram of calcined mineral and in average numbers of H_2O molecules per unit-cell.

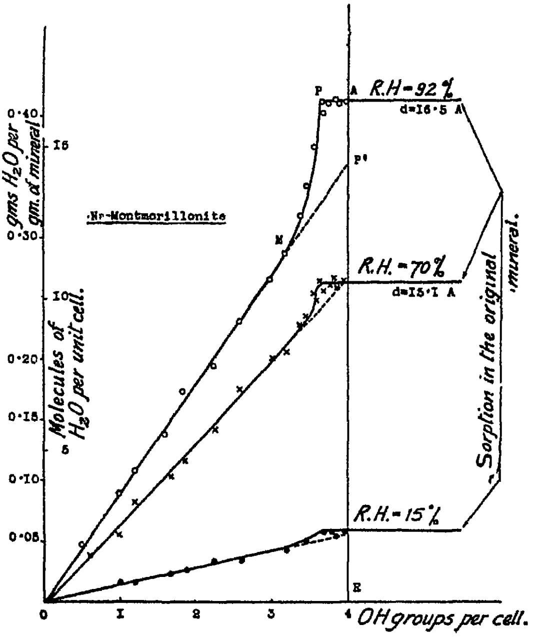


FIG. 2(*a*).

We note: (1) The alteration of Ca-montmorillonite begins with a steep drop *QP*. This discontinuity is accompanied by a characteristic transformation of the mineral which suddenly loses its property of disaggregation in water. This is not shown, however, by the X-ray diagrams, and it is therefore a superficial phenomenon. Thus the steep drop *PQ* is produced by the diminution of water of *adsorption* and it is brought about by the elimination of a small number of OH-groups on the surface of the particles.

(2) The curve *PNMO* of Ca-montmorillonite is accompanied by the transformations of X-ray diagrams described above. This is due to the immobilisation of an increasing number of sheets. The curves, which fall into three linear sections, show two distinct stages in the decrease of the hydratable surface. In the first stage (section *PN*, temperatures 350°-600° c.), the layers are bound together as a result of the partial elimination of their OH-groups. In the second (section *MO*, temperatures greater than 700° c.) the immobilisation of the layers is brought about by the total elimination of their OH-group. This begins at the point *N*, about 600° c., but since the total dehydration first affects the sheets already

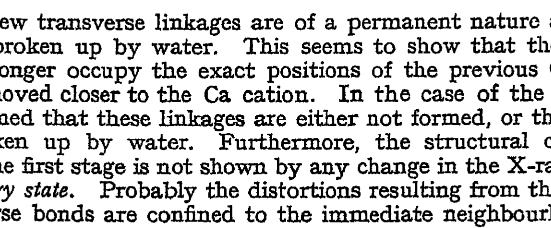
immobilised in great number by the first stage, the effect on the hydratable surface is very small. This initial part of the second stage is shown by section *NM*.

(3) For Na-montmorillonite a partial dehydration of the layers occurs at the beginning. The OH-groups eliminated by heat are more completely reconstituted than with Ca-montmorillonite, but are not totally reconstituted. This partial deficiency does not have any effect on the absorption capacity (section *AP*, temperatures 350°-550° c.). The decrease begins here suddenly at *P*, about 550° c. and rapidly assumes a linear aspect, the layers becoming fixed as a result of the total elimination of their OH-groups. Again, the Na-montmorillonite shows two stages of alteration, but the first stage has no fixing effect on the altered sheets. Here also the start of the diminution of sorption capacity is accompanied at *P* by the sudden disappearance of the capacity of the mineral to disaggregate and become plastic in water, but the superficial phenomenon cannot be separated as in the case of Ca-montmorillonite.

Probable Mechanism of Irreversible Dehydration.

(*a*) At Moderate Temperatures (350° C.-550° C. First Stage).—The fact that the elimination of part of the OH-groups immobilises the layers of Ca-montmorillonite but leaves unaffected the hydratable surface of Na-montmorillonite, indicates that cross linking is brought about by polyvalent exchangeable cations. On the other hand, the average number of OH-groups whose elimination brings about the bonding of a pair of sheets can be found easily from Fig. 2. It is represented by the length *SE*, which is intercepted on the abscissa by the prolongation of the line *PN*. This number is about 1·3 OH-groups per unit-cell. It represents the elimination of 0·65 molecules of water of constitution and the liberation of an equal number of oxygen atoms. This figure of 0·65 is striking since it is nearly equal to the base-exchange capacity of the mineral expressed in valencies per unit-cell, i.e. to double the number of cations which should occupy the spaces between the layers.

This coincidence is surely not fortuitous, and it shows that the liberation of 2 atoms of oxygen per Ca cation leads to the formation of oxide linkages —O—Ca—O— in the spaces between the layers affected by the partial dehydration, according to the following scheme:



These new transverse linkages are of a permanent nature and can no longer be broken up by water. This seems to show that the 2 oxygen atoms no longer occupy the exact positions of the previous OH-groups, but have moved closer to the Ca cation. In the case of the Na cation, it is presumed that these linkages are either not formed, or that they are easily broken up by water. Furthermore, the structural change produced in the first stage is not shown by any change in the X-ray diagrams taken *in dry state*. Probably the distortions resulting from the formation of transverse bonds are confined to the immediate neighbourhood of the Ca cations and only appreciably affect the small number of adjoining O-atoms, displacing only very little the neighbouring metallic atoms; the framework of the silicate thus retaining its rigidity.

(b) At High Temperatures (Greater than 550° C. Second Stage).—The fixing of the layers which accompanies their total dehydration appears to be independent of the nature of the exchangeable cations. X-ray diagrams taken *in the dry state* now reveal modifications which are more and more marked as the number of layers affected increases. These modifications are the same for Ca-montmorillonite and Na-montmorillonite. They can be summarised as follows.

(1) Half-way through the second stage (at about 700° c.), appears a phase of spinel structure, resulting from an advanced state of decomposition of the silicate. Grim and Bradley report the appearance of this phase at higher temperatures,[6] but it is probable that its formation follows closely the beginning of the second stage of dehydration.

(2) The altered silicate makes up the greater part of the total substance, but this principal phase retains only the rough structural scheme of the original clay. The diagram shows signs that destruction is setting in; the (*hk*) lines due to the two-dimensional structure of the layer vanish at large angles. In contrast with these lines, the fundamental (ool) reflections do not show any falling-off of intensity at large angles. Consequently, in dehydrated layers the hexagonal skeleton of the silicate must be partly broken up or disturbed, but the stratified structure of the mineral as a whole keeps a good average regularity.

(3) On orientated samples the relative intensities of the (ool) reflections become greatly modified. The new relative intensities are no longer compatible with the structural model of pyrophyllite. A better fit with the modified intensities can be obtained if a certain number of SiO_4 tetrahedra are now assumed to bind the layers together. This type of alteration seems to be a possible form of the beginning of decomposition of the silicate in two phases—the cubic oxide and silica. But this is tentatively suggested, for it is difficult to place much confidence in intensities of a phase so poorly organised.

We deduce from (1)-(3) that layers which have been definitely deprived of the whole of their OH-groups are appreciably decomposed. It appears unlikely that the spaces between these layers in course of destruction continue to be filled by the same cations as in the original clay; it is more probable that a certain number of polyvalent atoms Al, Mg spread outside the altered structure and that some of these atoms are retained in these spaces and then play the same part as the Ca cations during the first stage. Thus the case of Na-montmorillonite, oxide linkages can appear in this way, due to polyvalent ions being expelled from the sheets.

To conclude, OH-groups do not seem to play any active part in the mechanism of the fixation of water by clay: that layers of Na-montmorillonite can be deprived of part of their OH without losing their capacity for hydration, is significant in this respect. *The elimination of OH only suppresses the capacity for hydration inasmuch as the resulting structural modifications leads to the formation of linkages which cannot be broken up by water.*

In a clayey mass, water overcomes not only the forces of attraction between the layers inside a primary particle, but also those forces between one particle and another, to which are due the cohesion of the whole aggregate. The second of these effects leads first to disaggregation and then to dispersion.

The superficial forces existing between particles in an aggregate are generally attributed to the exchangeable cations outside these particles. They are consequently the same sort of forces as those acting between the elementary layers constituting the particles. Observations on irreversible dehydration confirm this. Thus the beginning of the decrease of the hydratable surface always coincides with the sudden disappearance of the phenomenon of disaggregation in water. Thus, bonding between the particles makes the aggregates non-dissociable by water and constitutes the "Baking" effect. "Baking" takes place at the beginning of the first stage of dehydration in Ca-montmorillonite, and at the beginning of the second stage in Na-montmorillonite, and at the beginning of the second stage in Na-montmorillonite. It therefore replaces the ionic bonds between particles by the same sort of oxide linkages as those formed between the elementary layers.

---

[6] *J. Amer. Chem. Soc.*, 1940, **23**, 242.

The principal interest of the analysis of irreversible dehydration would seem to lie less in indicating the probable mechanism of alteration than in the selective manner in which this alteration is manifested. The behaviour of montmorillonite towards water appears here to be due to three processes which can be differentiated by heat treatment. They are, in order of increasing sensitivity: (1) *hydration* of the spaces between the elementary sheets with formation of Hendricks' layers, (2) *adsorption* on the exterior of small organised domains which are the primary particles; (3) *plasticising* and dispersion.

The destruction of the adsorption capacity is shown in Ca-montmorillonite by the discontinuities *QP* of Fig. 2. Unbreakable linkages hold the particles together and make their surface inaccessible, as in the case of the sheets, but the effect on the surface of the particles is felt before that on the interior.

It will be seen, on the other hand, that in Na-montmorillonite no separate inhibiting of the adsorption capacity is obtained: the second stage of dehydration reduces the accessible surfaces without discriminating between the interior and the surface of the particles. It is, however, possible to heat with care Na-montmorillonite so as to bring it to the state represented by a point kust below point *P* (Fig. 2*a*). The clay then keeps its surfaces of hydration and adsorption intact, but is no longer either plasticised or disaggregated in water.

It thus appears that the plasticising action of water is not related either to the bulk state of the clay or to the surface area of the primary particles. Its extreme sensitivity means that the formation of a *very small* number of unbreakable bonds is enough to prevent it. These considerations will be used to determine the part played by water of hydration and water of adsorption in the swelling process.

The Isotherm of Hydration of Ca-Montmorillonite.—The water of hydration and adsorption in unaltered clay can be determined separately, using the curves in Fig. 2. It would be natural to take as a measure of the water of hydration the quantities measured by *EP*, the difference between the total quantity of water fixed reversibly and the discontinuity *PQ*, which we attributed to adsorption. The figures so obtained are still too high in R.H. $>$ 80 % in relation to the spacings deduced from X-ray diagrams. On the other hand, a fair agreement with X-ray data is obtained by taking for the water of hydration the length *EP′* intercepted by the continuation of the straight line *OM* which represents the second stage of dehydration. This choice amounts to admitting: (*a*) that the discontinuity *PQ* represents the disappearance of the greater part, but not of the whole, of the adsorption capacity; (*b*) that the remainder of this capacity is gradually lost along the contour *PNM*, the samples corresponding to the section *OM* (second stage) no longer containing anything but water of hydration. The second hypothesis is borne out by the tendency of Ca-montmorillonite to alter superficially at first. Taking the length *EP′* to be a measure of the water of hydration is the same as using the statement made above; i.e. that at every point of the segment *OM* the material is made up of a totally dehydrated and inactive portion, and a portion composed of sheets which retain all their OH-groups and are capable of hydration by moving away freely.

By this method the isotherms representing the water of hydration alone, may be traced. Fig. 3 (curve 1) gives the isotherm of hydration at 30° c. for Camp-Berteau Ca-montmorillonite. The point corresponding to a 100 % R.H. is calculated from the diffraction diagram given by clay under water.

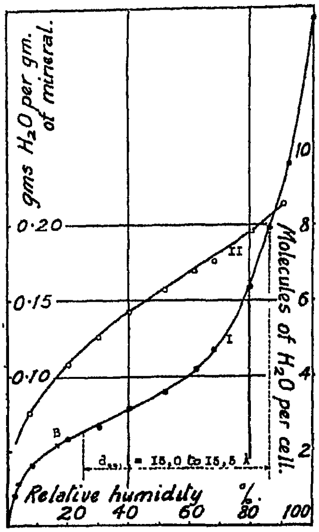


Fig. 3.—Ca-Montmorillonite.
I.—Hydration isotherm at 30° c.
II.—Adsorption isotherm at 30° c.

We note: (1) The spacing d between the sheets remains at 15·5 A. over a large range of R.H.'s. It increases from the point of the isotherm corresponding to the fixation of 8 H_2O molecules per unit-cell, i.e. equivalent to the formation of two complete Hendricks' layers.

(2) The ordinate of point B of curve 1 represents the first complete molecular "layer,"[7] or rather the number of molecules fixed with the highest heat of hydration.

This number is about $2H_2O$ molecules per unit-cell. The number of Ca cations, measured by the exchange capacity, is about 0·35. The quotient 2/0·35 of approximately 6 means that the first stage in hydration is the formation of groups of 6 H_2O molecules round every Ca cation, as deduced previously by Hendricks, Nelson and Alexander,[8] from the differential thermal analysis of montmorillonites saturated with alkaline earths.

(3) The above, as well as the wide range of sorption for which the X-ray diagram remain stationary (where $d = 15{\cdot}5$ A.), suggests that the "hydrate" with two layers of water begins to form immediately after the hydration of Ca cations, without passing through the one-layer stage, as with Na-montmorillonite and H-montmorillonite. The X-ray diagrams of the slightly hydrated states of Ca-montmorillonite confirm this: no basal reflections are observed as a rational series corresponding to $d_{001} = 12{\cdot}5$ A., whereas they are with H- and Na-montmorillonite.

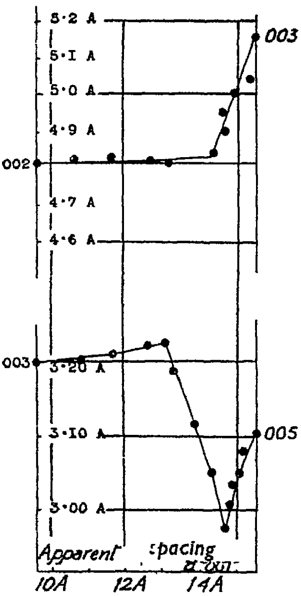


Fig. 4.—Ca-Montmorillonite.

In an irrational series an estimate of the actual spacings can be made by measuring a sufficient number of apparent spacings.[9] Fig. 4 shows (in Camp-Berteau Ca-montmorillonite) the variation of two apparent spacings of approximately 3 A. and 5 A. according to the variations of the fundamental apparent spacing, when this varies between 9·6 A. (in the dry state) and 15·5 A. (two-layer "hydrate"). These curves can be interpreted by the simultaneous presence of a succession of 3 fundamental spacings of 9·6 A., about 14 A., and 15·5 A.; in the right-hand portion of the two curves, the "hydrates" of 14 A. and 15·5 A. alone are present. The spacing of 14 A. is due to the hydration of the Ca cations and it seems likely that the H_2O molecules fixed by each cation surround it octahedrally.

---

[7] Brunauer and Emmett, *J. Amer. Chem. Soc.*, 1935, **57**, 1754, and 1937, **59**, 2682.
[8] *Ibid.*, 1940, **62**, 1457.
[9] Hendricks and Teller, *J. Chem. Physics*, 1942, **10**, 147.

To sum up, the hydration of Ca-montmorillonite takes place by: (*a*) formation of $Ca(H_2O)_6$ groups and a resulting spacing of about 14 A.; (*b*) filling of the remaining space, and increase of the spacing to 15·5 A. The sheets retain this constant spacing until two complete Hendricks' layers have been formed; (*c*) formation of a third, and then of a fourth complete molecular layer, increasing the spacing by 3 A. per layer.

Water of Hydration in Na-Montmorillonite.—The method of separation used above is useless for Na-montmorillonite. Firstly, no discontinuity attributable to water of adsorption appears in Fig. 2*a*. Secondly, the length *EP'*, which is intercepted by the continuation of the line *OM*, gives figures which are too high by comparison with the possible contents deduced from the spacings shown by X-rays. It seems unlikely that the layers of water are more compact than in Hendricks' scheme—e.g. the density measurements disprove this—and it must therefore be concluded that the lengths *EP'* represent the sum of the water of hydration and adsorption, and that the second stage of alteration reduces without discrimination all surfaces accessible to water. It may be noted that the segment *EP'* is equal to the whole sorption of the original clay, except in R.H. > 90 % (see Fig. 2*a*). The difference observed in the curve at 92 % R.H. may be associated with the onset of plasticity, which becomes perceptible above about 90 % R.H.

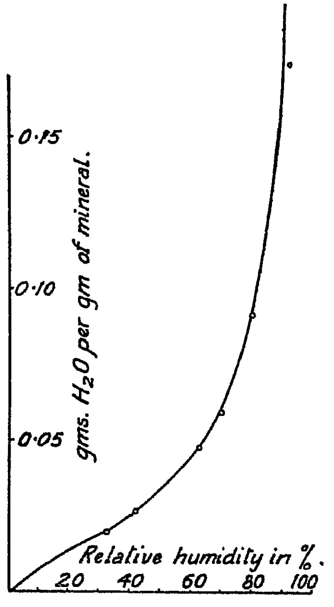


FIG. 5.—Na-Montmorillonite. Adsorption isotherm at 30° c.

An estimate of the water of hydration can only be made in Na-montmorillonite from the X-ray diagrams, and these do not show a stationary region comparable to that of Ca-montmorillonite; thus the hydration seems to occur by *complete* molecular layers. However, a more or less accurate measurement of the average spacing can only be made if the basal reflections are sufficiently sharp and intense, e.g. with Na-montmorillonite in the region of high humidities and in a narrow range of low humidities, and consequently only a few points can be plotted.

Water of Adsorption.—The isotherm of adsorption of Ca-montmorillonite is given by the difference between the ordinates of the whole isotherm and those of the isotherm of hydration (curve II of Fig. 3); thus adsorption is considerable even in low humidities.

For *Na-montmorillonite:* (*a*) In 33 % R.H. (Camp-Berteau mineral) the X-ray diagram shows the formation of a homogeneous "one-layer hydrate". The total water fixed is 0, 12 g. per g. of heated matter, i.e. 20 % more than the quantity needed for hydration with one layer.

(*b*) The general isotherm of the same Na-montmorillonite obeys the isotherm equation of Brunauer, Emmett and Teller [10] up to a relative humidity of 35 %, viz.,

$$h = h_m \frac{a}{(1-a)} \frac{c}{(1+(c-1)a)} \quad . \quad . \quad . \quad (1)$$

where h = the total mass absorbed, h_m = the mass of the first complete molecular layer, a = the relative pressure (on R.H. here), and c approximately = exp. $\Delta W/RT$ with ΔW = the difference between the heat of fixation of the first layer and the heat of liquefaction. The application of this equation gives $\Delta W = 815$ cal./mole., and $h_m = 0{\cdot}123$ g./g. of calcinated matter.

---

[10] *J. Amer. Chem. Soc.*, 1938, **60**, 309.

(*c*) From (*a*) and (*b*) it can be deduced that the free surface liable to adsorption represents 20 % of the hydratable surface of the silicate sheets, and the adsorption isotherm can be calculated from equation (1), if $h_m = 0{\cdot}020$ g. and c keeps the value found above (cf. Fig. 5). Except in very high R.H., the "experimental points" resulting from the difference between the total sorption and the water of hydration calculated from the average spacing are found to lie on the theoretical curve.

Therefore the adsorption of water on the exterior of the primary particles takes place in Na-montmorillonite in the form of multilayers, with a heat of condensation close to the heat of liquefaction for every layer except the first.[10] Moreover, where the total surface of the sheets is about 4×10^6 sq. cm. per g. the free surface must be about $0{\cdot}8 \times 10^6$ sq. cm./g. If it is assumed that the water adsorbed is fixed on the (001) faces of the primary particles, double the above figure must be taken for the free surface, i.e. $1{\cdot}6 \times 10^6$ sq. cm./g. The average number of silicate sheets in a primary particle can thus be determined as about 5. The corresponding figure for Black Hill montmorillonite is about 15. Returning to the isotherm of Ca-montmorillonite, we see that the quantities of water adsorbed, as well as the shape of the isotherm, are incompatible with an adsorption of the multilayer type. Bulk condensation seems to take place in the spaces between the primary particles. Hydrated Ca cations must play an important part in this phenomenon.

Fundamental Differences between the Phenomena of Adsorption and Hydration.—A parallelism between adsorption and hydration exists inasmuch as the water breaks down the ionic bonds inside the particles as well as between aggregated neighbouring particles. But there are several fundamental differences. It is possible that in Na-montmorillonite the first molecular layer of water adsorbed possesses the hexagonal structure of the water of hydration. But it is improbable that the "organising" action of the silicate can be felt at any great distance. With increasing distance from the silicate surface the configuration of water molecules approximates more nearly to that of the liquid state. The quantity of water in the plastic gels of Na-montmorillonite is large enough to produce characteristic halos to monochromatic X-ray diagrams; the photometry of these diagrams shows that the greater part of the water in the gel has the structure of liquid water.

Secondly, the distance between the silicate does not increase indefinitely, whereas the increasing separation of the particles finishes by their dispersion in water. The non-dispersion of the elementary sheets is probably to be associated with their two-dimensional form and high degree of parallelism, these factors reducing the variation of the cohesive forces with distance. Between neighbouring particles there is no such complete parallelism, and mutual cohesion is both weaker and more sensitive to increase of distance. It might be expected that this faculty of dispersion would be modified by mutual orientation of the particles, and this, in fact, is observed.

Reversibility of the Phenomenon of Dispersion. Spontaneous Orientation.—In drying a suspension of Na-montmorillonite (of depth sufficient to eliminate the action of the underlying surface) the small mobility of the particles and more particularly the difference in their sizes allows the cohesion forces to immobilise them long before the disorder produced by the Brownian movement has disappeared. Disorientated aggregates are established and can be dispersed again in water. The dispersion can be said to be reversible if, in the course of alternate suspension and drying, there is found in each new suspension the same particle size distribution (by centrifuging, for example). In Na-montmorillonite this reversibility exists for the unfractionated sample and for coarse fractions, but not for those which are very fine and close to the monodispersed state. A fine fraction dried as above gives, instead of the usual split blocks, a mass with the appearance of crumpled sheets of paper. X-rays reveal in the flaky substance a pronounced parallelism of the (001) faces of the particles. This is a case of particle-to-particle orientation produced in the absence of any exterior action. It is facilitated by the mobility of the small particles and also by a uniform size distribution. If the orientated mass is put into water again, it swells, but is not dispersed.

It appears then that the prerequisite for reversibility of dispersion is the presence of enough large-sized particles to prevent the mutual orientation of the fine particles. This is confirmed by mixing in varying proportions two different fractions in suspension, e.g. a fraction $A > 0{\cdot}1\mu$ and one $B < 0{\cdot}03\mu$. As the proportion of A increases the spontaneous orientation during drying diminishes, and the quantity of B which passes into suspension when water is added increases. If there is sufficient A (about 50 % of the mixture) no further spontaneous orientation is observed and the total reversibility of dispersion and aggregation by drying is recovered.

Stable Orientated Aggregates.—This name is given to the aggregates of fine fractions described above ; they can more conveniently be formed by evaporating a thin layer of the suspension on a flat surface . in this way the mutual orientation of neighbouring particles is not increased, but instead of a tangled mass, a single orientated film is obtained. Orientated films can also be obtained from any well-dispersed non-fractionated sample.[11] X-rays show no difference in the degree of orientation in the two cases, but in the latter the deposition surface alone imposes an average preferential orientation on the whole of the mass. With very fine fractions the interaction between neighbouring particles is shown by the tendency of the deposited films to become creased during formation.

Both types of film swell in one dimension only, i.e. the thickness increases. But their behaviour in water is different, viz. : (*a*) in distilled, undisturbed water, the orientated film of the non-fractionated material disaggregates in a short while in the same way as an unorientated aggregate ; (*b*) under the same conditions a film obtained from a very fine fraction swells without being dispersed and keeps its definite outline ; its thickness increases about 20-fold. The X-ray diagram shows that this swelling has little effect on the average degree of orientation of the particles.

Stable orientated aggregates can also be obtained with Ca-montmorillonite which is hard to disperse and cannot be divided directly into fractions. Fractions of Na-montmorillonite in suspension are converted to the acid clay ; each fraction is then treated with a quantity of $Ca(OH)_2$ which is exactly equivalent to its " base exchange " capacity. By avoiding excess lime, flocculation is prevented [12] and well-dispersed suspensions of Ca-montmorillonite can be obtained. Spontaneous orientation of fine fractions ($< 0{\cdot}05\mu$) occurs under the same conditions as in Na-montmorillonite. These films resist the disintegration action of distilled water. Moreover, no apparent swelling is observed and the films are not perceptibly plasticised. To summarise, the mutual orientation of fine particles prevents dispersion of Na-montmorillonite, and inhibits swelling and plasticising in Ca-montmorillonite.

Water of Adsorption as a Function of the Variations of Apparent Particle Size and the Degree of Organisation of the Aggregates.—An estimate may be made of the water of adsorption in different fractions of Na-montmorillonite and Ca-montmorillonite, firstly in the state of unorinetated aggregates (blocks and powder), and secondly in the state of orientated films.

Fine fractions of dry Na-montmorillonite cannot be collected in an unorientated form, and must therefore be examined when mixed with an equal mass of the fraction $> 0{\cdot}1\ \mu$. Fine fractions of Ca-montmorillonite are easily collected in the unorientated state by flocculating the suspensions with excess $CaCl_2$, which is later eliminated by washing. Observations made in different R.H.'s < 95 % show that the proportion of water of adsorption (as well as that of water of hydration) is always the same, and in consequence it is not dependent either on the apparent size of the particles or their mutual orientation. The free surface liable to adsorption is therefore the same in all fractions. Thus, in a dispersion of Na-montmorillonite the particles are not "monocrystalline" but, at least in large and medium sized fractions, stable associations of primary particles in varying numbers. They are stable micro-aggregates, probably of the same kind as the macroscopic aggregates which resist dispersion. This distinction between the primary particle and the multiple particles liberated by dispersion accords with the X-ray diagrams which are practically identical for all fractions. The primary particles are probably in a free state only in the finest fractions ($< 0{\cdot}03\ \mu$).

---

[11] Clark, Grim, Bradley, *Z. Krist.*, 1937, **96**, 322 ; Nagelschmidt, *J. Sci. Instr.*, 1941, **18**, 5.

[12] Glaeser, *Compt. Rend.*, 1946, **222**, 1179-81.

Swelling of Stable Orientated Aggregates. *Na-Montmorillonite :* Fairly homogeneous films having an approximately constant thickness can be prepared on about 1 sq. cm. of the sample. These films are convenient for the comparative study of apparent swelling (variation of the film thickness) and the internal swelling of the primary particles (variation of the average spacing between elementary sheets). For Na-montmorillonite the observations are summarised in the following table :

	Relative Humidity (R. H.).	Total Absorption in g. of Water per g. of Dry Substance.	Average Spacing (d) in A.	Hydration Water (Average Number of Layers).	Water on the Exterior of the Primary Particles in g./g. of Dry Substance.	Apparent Swelling.
I.	0	0	9·6 A.	0	0	No swelling
	↓	↓	↓	↓	↓	↓
	90 %	0·40	16·2 A.	2·2	0·18	Beginning of swelling and plastification
	↓	↓	↓ The intensity of the (001) reflection increases with swelling.	↓	↓	
II.	96 %	0·60	18·5 A.	3	0·30	About 30 % swelling
	↓	↓	↓	↓	↓	↓
	99 %	1·00	19·5 A.	3·3	0·80	About 100 % swelling
III.	Absorption in contact with water	↓ About 5 ↓	The apparent spacing remains fixed at 19·5-20 A. but the (ool) reflections gradually vanish Complete disappearance of (ool) reflections	↓	↓	↓ Swelling increases to about 20 times the original thickness ↓

The reversible hydration of the aggregate can be seen to pass through three successive stages.

STAGE I.—No swelling is produced in spite of the absorption of 0·40 g. of water. X-ray diagrams, however, indicate that the volume of the primary particles is increased by 70 %. Two points should be noted : (*a*) the apparent density of the dry aggregate was about 0·4 of the theoretical density of the dry mineral, i.e. the texture of the aggregate has

60 % of empty spaces, which permits the fixation of 0·55 g. of water without any variation in volume of the whole. (*b*) The texture of the aggregates must be such that the voids *absorb* the increase in volume resulting from the increase of the spacings between the silicate sheets. This absorption of the internal swelling is only possible if the primary particles are connected laterally, with no contact between their (001) faces. The orientated aggregate must be considered as a succession of superimposed sheets of primary particles joined together laterally, with few contacts between successive sheets. The average width of the spaces between the sheets is of the same order as the average thickness of the primary particles.

STAGE II is manifested by the beginning of the plasticising and swelling of the film. We note : (*a*) the increase of the spacing between the silicate sheets is relatively small and not in proportion to the swelling observed ; (*b*) the amount of water outside the primary particles becomes much larger and finally preponderates. The swelling of the aggregate must therefore be ascribed to the increase in the spaces between the sheets of particles. It appears that as soon as the original spaces are filled, the water dissociates the few connections between neighbouring particle sheets and forces the latter further apart ; it is this which must be responsible for the plasticising of the aggregate.

STAGE III is not characterised at the outset by any peculiarity of behaviour in the apparent swelling. On the X-ray diagrams it is marked by the progressive fading of the (ool) reflections, which finally disappear altogether. The following facts show that this phenomenon is not to be interpreted in terms of a large increase of the spacing between the elementary silicate sheets : (*a*) While the intensity of (ool) interferences diminishes, the apparent spacing remains fixed at 20 A., indicating some disorder in the primary particles rather than their expansion. This disorder in the highly swollen state may be either lack of parallelism of the layers or non-uniformity of the inter-layer spacings. (*b*) It seems legitimate to extrapolate from the above interpretation of Stage II and to assume that the water outside the particles forms an increasingly preponderant part of the total water. (*c*) The aggregate can be heated with care until it reaches the state represented by point *P* on Fig. 2*a* (cf. above). By this treatment plasticising and swelling are prevented while the hydration and adsorption surfaces remain intact. In R.H. < 90 % there is no distinction between the diffraction diagrams of the aggregate when so treated and the original aggregate. If placed in water, the treated aggregate produces intense (ool) reflections and shows an average spacing of 20 A. This indicates that swelling can be prevented by the formation of a relatively small number of indissoluble connecting links between the sheets of particles, the particles themselves retaining their freedom of expansion. Thus it appears that swelling and plasticising are due to the formation of large spaces between the particles rather than to increasing separation of the elementary silicate sheets. For particles containing an average of 10 silicate sheets the mean interparticulate distance at the limit of swelling is of the order of 2000 A.

Ca-Montmorillonite.—It has been shown that in the case of Ca-montmorillonite only stage I exists. Their presence of Ca ions is therefore sufficient to prevent the complete dissociation of the connecting links between the sheets of particles. The fundamental spacing *d* reaches about 20 A. in water, and the (ool) reflections are intense. The apparent density of the films is the same as for Na-montmorillonite.

Summary.

(1) The study of the decrease of the absorption capacity of montmorillonite when subjected to the progressive elimination of the OH-groups shows that the latter do not play an active part in the mechanism of water

fixation. Their elimination reduces the absorption capacity inasmuch as it brings about the formation of indissoluble linkages between the elementary sheets and between the particles. These linkages are formed by means of Ca cations in Ca-montmorillonite. In Na-montmorillonite, they seem to be due to polyvalent atoms expelled during incipient decomposition of the silicate.

(2) In Ca-montmorillonite, the change takes place first at the surface of the particles. This allows a distinction to be made between the water of hydration (fixed between the elementary sheets) and the water of adsorption (fixed between the particles). The isotherm of hydration drawn in this way agrees with the indications of the X-ray diagrams.

(3) The observations on adsorption show that various levels of organisation exist in montmorillonite : (*a*) Primary particles consisting of stacks of parallel elementary sheets with an average of about 10 sheets per particle. (*b*) Stable micro-aggregates formed by the association of several primary particles which are nearly parallel and joined together laterally. (*c*) Aggregates comprising a large number of primary particles and micro-aggregates.

(4) The stable micro-aggregates are not destroyed when Na-montmorillonite is dispersed in water. Their structure is suggested by the properties of stable macroscopic aggregates which are formed spontaneously when a suspension of a very fine fraction ($< 0{\cdot}05\ \mu$) is dried. These aggregates swell in water without dispersing. The constituent particles show strong mutual orientation.

(5) An examination of the hydration of the stable orientated aggregates indicates that the swelling and plasticising of Na-montmorillonite are due to increasing inter-particulate spaces rather than to expansion of the individual primary particles.

Laboratoire Central des Services
Chimiques de l'Etat,
Paris.

IONIC FORCES IN THICK FILMS OF LIQUID BETWEEN CHARGED SURFACES.

By R. K. Schofield.

Received 29*th July,* 1946, *and as revised* 10*th October,* 1946.

In this paper an attempt is made to elucidate a type of swelling that is encountered in many systems where solid surfaces are separated by films of liquid. The surfaces may be those of separate particles or of fibrils forming a connected network. Soils and clays are the outstanding natural objects in the first category which also embraces a host of man-made pastes. In these systems it is important to distinguish broadly between irreversible shrinkage due to rearrangement of the particles, and reversible shrinkage associated with the dependence of film thickness on pressure. Natural fibres provide examples in the second category. The chief complication here is that the solid matter of such a system may itself swell by molecular sorption.

It is only when the solid matter of these systems is very finely divided, that considerable swelling and shrinkage can occur as a result of alterations in film thickness, and it is then very difficult to make accurate measurements of particle size or of surface area. X-ray analysis has been successfully applied to the measurement of the swelling of the clay mineral, montmorillonite (bentonite). Here it appears reasonable to regard each crystal as a stack of "particles," 10 A. thick, separated by "films" which increase in thickness to about 7 A. at 90 % R.H., the exact value depending on the nature of the exchangeable ions.[1] In a natural clay subsoil the release of water to plant roots is essentially a reversible shrinkage caused by a reduction in the effective relative humidity (i.e. after allowing for the vapour pressure reduction due to salts) from roughly 99·99 % to 99 %. Bradley, Grim and Clarke[2] have recorded a basal spacing of 21·4 A., i.e. a "film" thickness of 11 to 12 A. in very wet montmorillonite, but no extensive study has yet been made of thickness with different cations at relative humidities close to 100 %.

In these circumstances the author has made a different approach to the problem. The relationship between film thickness and swelling pressure must be governed by ascertainable laws. If a quantitative relationship having a satisfactory theoretical basis can be verified under favourable experimental conditions, it may be possible to invert the attack on more complex systems. Thus it might be possible from measurements of film pressure, or the corresponding vapour pressure lowering, to make an estimate of film thickness so much closer than can be obtained in any other way, as to provide the best estimate of surface area.

The distribution of ions in a solution in contact with a charged surface was first discussed by M. Gouy,[3] and further elucidated by D. L. Chapman.[4] The main object of these studies was to obtain the relationship between the surface density of charge and the potential across the diffuse component of the electric double layer. H. Muller[5] obtained this relationship for the general case where the ions of the electrolyte have unequal charges. The repulsion between two charged surfaces arising from the osmotic activity of the ions in the intervening film has been considered by I. Langmuir,[6] and by A. Frumkin and A. Gorodetzkaja,[7] who obtained approximate relationships.

In the following treatment, consideration is first given to the film on a single surface. It emerges that under the conditions postulated the film separating two similar parallel surfaces should be twice as thick. This is not a thermodynamic necessity, and is probably not strictly true, but it should be practically true for thick films, and a good first approximation for thin ones.

Let us suppose we have a liquid film of thickness X adhering to a solid surface, on which there are situated electrically charged atom-groups at a surface density Γ. We will also suppose that the ions in the film forming the diffuse component of the electric double layer are all of one kind having, consequently, a sign opposite to that of the charges on the surface, and having valency ν. Let the ionic concentration at a distance x from the outer surface of the film (i.e. $X - x$ from the solid surface) be c, and let that at the outer surface ($x = 0$) be c_0. We will further suppose that c is small enough for the ions to be in "ideal" solution. Then the electrical potential at x will differ from that at the outer surface by $\frac{RT}{\nu F} \ln \frac{c}{c_0}$. The electric intensity at x is therefore given by $\frac{RT}{\nu F} \cdot \frac{1}{c} \cdot \frac{dc}{dx}$. This electric intensity may be considered to arise from surface densities of charge $\frac{\epsilon}{4\pi} \cdot \frac{RT}{\nu F} \cdot \frac{1}{c} \cdot \frac{dc}{dx}$ and $-\frac{\epsilon}{4\pi} \cdot \frac{RT}{\nu F} \cdot \frac{1}{c} \cdot \frac{dc}{dx}$ separated by an imaginary plane distant x from the outer surface. The stress across this plane due to electrostatic attraction will thus be $\frac{\epsilon}{4\pi}\cdot\frac{RT}{\nu F}\cdot\frac{1}{c}\cdot\frac{dc}{dx}\times\frac{1}{2}\frac{RT}{\nu F}\frac{1}{c}\frac{dc}{dx}$. Again assuming the ions to be in ideal solution, we may equate this stress to the amount by which the van't Hoff pressure, RTc, at x exceeds the van't Hoff pressure, RTc_0, at the outer surface, where dc/dx is evidently zero.

---

[1] Hendricks, Nelson and Alexander, *J. Amer. Chem. Soc.*, 1940, **62**, 1457.

[2] *Z. Krist.*, 1937, **97**, 216.

[3] *J. Physique*, 1910, **9**, 347.

[4] *Phil. Mag.*, 1913, **25**, 475.

[5] *Kolloidchem. Beihefte*, 1928, **26**, 257.

[6] *Science*, 1938, 88, 430, and *J. Chem. Physics*, 1938, **6**, 873.

[7] *Acta Physicochim.*, 1938, **9**, 327.

Writing $\beta = \frac{8\pi F^2}{\epsilon RT} = \frac{8{\cdot}48}{\epsilon} \times 10^{16}$ cm./millimoles at 25° c., we have

$$\left(\frac{1}{c}\cdot\frac{dc}{dx}\right)^2 = \nu^2\beta(c - c_0). \quad . \quad . \quad . \quad . \quad (1)$$

The author has been furnished by Mr. M. H. Quenouille with the solution to this differential equation which is

$$c = c_0 \sec^2 \frac{\nu x\sqrt{\beta c_0}}{2}. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Evidently $\Gamma = \nu\int_0^X c\, dx$.

Substituting for c, integrating and rearranging, we have

$$X = \frac{2}{\nu\sqrt{\beta c_0}} \cot^{-1} \frac{2\sqrt{c_0}}{\Gamma\sqrt{\beta}} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Clearly as $\Gamma \to \infty$, $X \to X_{\text{max.}}$, where

$$X_{\text{max.}} = \frac{\pi}{\nu\sqrt{\beta c_0}}. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

For large values of $\Gamma\sqrt{\beta}/2\sqrt{c_0}$, we have

$$X_{\text{max.}} - X \doteqdot \frac{4}{\nu\beta\Gamma}. \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Writing p_0 for the Van't Hoff pressure at the outer surface we have, so long as c_0 is very small,

$$p_0 = RTc_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

p_0 is then the excess of the hydrostatic pressure at the outer surface of the film over that of bulk liquid in equilibrium with it at the same horizontal level. In considering the application of the above treatment to a film between two similar parallel surfaces it will be seen that equation (2) can be used with both positive and negative values of x. p_0 is therefore the swelling pressure between two such surfaces separated by a film of thickness $2X$ under the conditions postulated. The equation given by Langmuir [6] can be obtained by eliminating c_0 between equations (4) and (6).

B. Deryagin and M. Kussakov [8] have measured by optical interference the thickness of films of liquid between solid surfaces and bubbles of hydrogen. The corresponding values of p_0 were obtained from the surface tensions and curvatures of the free liquid/gas surfaces. They consider that, in general, the presence of these films cannot be attributed to the ions forming the diffuse component of an electric double layer, although they concede that these ions may contribute to the effect in the case of water films on mica and glass. Although Deryagin [9] quotes Langmuir, he does not, for these cases, present a comparison of observed and calculated values. These comparisons are made in Tables I and II.

The crystal structure of mica provides one negatively charged atom-group consisting of an aluminium ion in four co-ordination per 47 A.^{2}. The corresponding value of Γ is $3{\cdot}5 \times 10^{-7}$ milliequivalent/cm.^{2}. Taking $\epsilon = 80$ and $\nu = 1$, since there will be mainly potassium ions in the film together with some hydrions, we find from equation (5) that $X_{\text{max.}} - X$ is only 0·1 mμ. Hence we need not distinguish between X and $X_{\text{max.}}$.

---

[8] *Acta Physicochim.*, 1939, **10**, 25 and 153.

[9] *Trans. Faraday Soc.*, 1940, **36**, 203.

Taking the measured values of p_0, c_0 was calculated by equation (6), and then $X_{\text{max.}}$ was obtained by equation (4). The agreement with the observed thicknesses exhibited in Table I is very satisfactory. The smallness of c_0 justifies the assumption of ideal solution. In the case of water on glass, we do not know Γ for the glass surface, but it could be one-tenth that for mica without causing a measurable decrease in film thickness. $X_{\text{max.}}$ for water on glass was, therefore, computed in the same way as for water on mica. Three of the comparisons in Table II. are within the 10 mμ which Deryagin and Kussakov estimate to be the limit of accuracy of the optical measurements. In the other three comparisons the observed thicknesses are less than that calculated.

TABLE I

WATER ON MICA $\epsilon = 80$, $\nu = 1$.

p_0 (obs.), dynes/cm.^{2}	c_0 (calc.), moles/l.	$X_{\text{max.}}$ (calc.), mμ.	Film Thickness (obs.), mμ.
490	$0{\cdot}20 \times 10^{-4}$	216	195
770	$0{\cdot}31 \times 10^{-4}$	174	179
990	$0{\cdot}40 \times 10^{-4}$	153	165
1,860	$0{\cdot}75 \times 10^{-4}$	111	100
2,580	$1{\cdot}04 \times 10^{-4}$	94	90
6,050	$2{\cdot}44 \times 10^{-4}$	62	46
13,770	$5{\cdot}55 \times 10^{-4}$	41	30

In a later series of measurements, Deryagin, Kussakov and Tityevskaya[9] used solutions of sodium chloride against freshly blown lead glass. Mr. M. H. Quenouille has worked out the values of $X_{\text{max.}}$ for a univalent electrolyte. Space does not permit the reproduction of the mathematics, but the nature of the calculation will be evident from the headings of the columns of Table III.* The experimental values of p_0 cover the range 400 to 1,200 dynes/cm.^{2} The thicknesses given for $p_0 = 1{,}000$ dynes/cm.^{2} were obtained by interpolation between plotted points; m stands for the molarity of the sodium chloride solution.

TABLE II.

WATER ON GLASS $\epsilon = 80$, $\nu = 1$.

p_0 (obs.), dynes/cm.^{2}	c_0 (calc.), moles/l.	$X_{\text{max.}}$ (calc.), mμ.	Film Thickness (obs.), mμ.
530	$0{\cdot}21 \times 10^{-4}$	211	151
1,260	$0{\cdot}51 \times 10^{-4}$	135	130
1,490	$0{\cdot}60 \times 10^{-4}$	125	129
3,440	$1{\cdot}39 \times 10^{-4}$	82	71
5,830	$2{\cdot}35 \times 10^{-4}$	63	21
5,910	$2{\cdot}38 \times 10^{-4}$	63	40

Values calculated by Deryagin,[9] using an approximate formula due to Frumkin and Gorodetzkaja,[7] are of the wrong order of magnitude owing to the accidental omission of a factor ϵ. Actually this formula, when rearranged and expressed by the symbols here used, can be written

$$X_{\text{max.}} \doteqdot \frac{1}{2\sqrt{\beta m}} \ln \frac{64RTm}{p_0} = \frac{\ln 4 \tan \theta_0}{\sqrt{\beta m}}.$$

When $p_0/4RTm$ is 0·1 or less, this equation reproduces the values of X_{max} given in Table III, but it is not applicable when $p_0/4RTm$ is as high as 100. The figure 140, in parenthesis, placed between concentrations 10^{-7} and 10^{-6}, to represent distilled water, is interpolated from Table II.

The lack of agreement between the calculated values of $X_{\text{max.}}$ and the observed thicknesses is puzzling. It is very strange that the theory which gives such good agreement in the case of distilled water should break down when as little as 10^{-4} molar sodium chloride is present. No satisfactory explanation can be offered at the present time.

---

* The complete elliptic function $F_1(\theta_0)$ is sometimes written $K(\theta_0)$.

Deryagin reports that with solutions containing polyvalent cations the films are too thin to be measured, and frequently form a contact angle of 1-3°. Here the departure from the simple theory is in the opposite sense. As this effect may be the counterpart for a film on one surface of the phenomenon of flocculation in colloidal systems its further study is likely to be particularly fruitful.

In discussing their results for hexane on steel, Deryagin and Kussakov assumed that ions could play no part, and concluded that long-range forces of another kind must be in operation. Some carboxylic acid anions could, however, have been present in the oil film, the corresponding hydrions having combined with the oxide layer on the steel surface giving it the corresponding positive charge. It is very difficult to rid hydrocarbon completely of traces of carboxylic acids. The values of $X_{\text{max.}}$ given in Table IV were obtained by taking $\epsilon = 2$ and $\nu = 1$. The agreement with the measured thicknesses is remarkably good. Fortunately Deryagin and Kussakov carried out a second set of measurements with 0·001 molar

TABLE III.

SODIUM CHLORIDE SOLUTIONS ON LEAD GLASS.

$\epsilon = 80$, p_0 (obs.) = 1,000 dynes/cm.^{2}

Sodium Chloride Conc. moles/1.	$\cot^2\theta_0 = \frac{p_0}{4RTm}$.	Xmax. (calc.) $= \frac{\sin\theta_0 . F_1(\theta_0)}{\sqrt{\beta m}}$. $m\mu$.	Film Thickness (obs.), $m\mu$.
10^{-7}	100·84	150	(140)
10^{-6}	10·084	146	
10^{-5}	1·0084	126	
10^{-4}	0·10084	77	122
10^{-3}	0·01008	36	74
10^{-2}	0·00101	15	49
10^{-1}	0·00010	6	48

TABLE IV.

HEXANE ON STEEL $\epsilon = 2$, $\nu = 1$.

Caprylic Acid Added moles/l.	p_0(obs.) dynes/cm.^{2}	c_0(calc.) moles/l.	Xmax. (calc.) $m\mu$.	Film Thickness (obs.) $m\mu$.
none	95	$0{\cdot}38 \times 10^{-5}$	78	64
,,	110	$0{\cdot}44 \times 10^{-5}$	73	67
,,	127	$0{\cdot}51 \times 10^{-5}$	68	68
,,	135	$0{\cdot}54 \times 10^{-5}$	66	53
,,	185	$0{\cdot}75 \times 10^{-5}$	56	54
,,	250	$1{\cdot}01 \times 10^{-5}$	48	50
,,	290	$1{\cdot}17 \times 10^{-5}$	45	47
,,	380	$1{\cdot}53 \times 10^{-5}$	39	42
0·001	107	$0{\cdot}43 \times 10^{-5}$	74	84
0·001	125	$0{\cdot}50 \times 10^{-5}$	68	80
0·001	160	$0{\cdot}65 \times 10^{-5}$	60	69
0·001	190	$0{\cdot}77 \times 10^{-5}$	55	66
0·001	235	$0{\cdot}95 \times 10^{-5}$	50	66
0·001	340	$1{\cdot}37 \times 10^{-5}$	41	57

n-caprylic acid added to the hexane. This addition caused an average of 15 mμ in the film thickness. This difference may well be significant, even if the absolute values are uncertain to 10 mμ. It seems permissible, therefore, to postulate that it was only with the addition of the caprylic acid that Γ was large enough for X to be practically equal to $X_{\text{max.}}$ and that, without this addition, $X_{\text{max}} - X$ averaged 15 mμ. Applying equation (5) we find $0{\cdot}63 \times 10^{-10}$ milliequivalent/cm.^{2} for Γ, which means one positive charge per 260,000 A^2.

The results for Vaseline oil on steel are set out in Table V. At 480 dynes/cm.^{2}, without the addition of oleic acid, the agreement is nearly as good as with hexane, but the addition of oleic acid seems to have a disproportionate effect. Indeed, the observed thickness at 180 dynes/cm.^{2} with 1 % of oleic acid added is greater than the calculated thickness for water. Vaseline oil is too indefinite a material to be suitable for such an experiment, and the addition of oleic acid may have influenced both the dielectric constant and the surface tension which is needed to obtain p_0. There is also the possibility that the reflexion coefficient of the steel surface was influenced by the presence of an adsorbed layer. These points would have to be checked before such an experiment could be accepted as proof of the existence of thick films that cannot be attributed to ions.

TABLE V.

VASELINE OIL ON STEEL $\epsilon = 2$, $\nu = 1$.

Oleic Acid Added moles/*l*.	p_0 (obs.), dynes/cm.^{2}	c_0 calc moles/l.	$X_{\text{max.}}$ (calc), mμ.	Film Thickness (obs), mμ.
none	250	$1{\cdot}01 \times 10^{-5}$	48	97
,,	480	$1{\cdot}94 \times 10^{-5}$	35	47
0·04	180	$0{\cdot}72 \times 10^{-5}$	57	403
0·04	490	$1{\cdot}98 \times 10^{-5}$	34	165
0·04	1220	$4{\cdot}90 \times 10^{-5}$	22	46

Deryagin's data [9] for *n*-hexyl alcohol on steel exhibit film thicknesses from 34 to 44 mμ less than the values of $X_{\text{max.}}$ calculated by taking $\epsilon = 16$ and $\nu = 1$. If we accept the experimental figures as measures of X, we find from equation (5) that Γ is approximately 2×10^{-10} milliequivalent/cm.^{2}, or one unit charge per 83,000 A^2. We may also put $x = X$ in equation (2) and use it to calculate c_X, the ion concentration when $x = X$. This works out about 5×10^{-5} mole/*l*. It may well be that the number of cations (presumably sodium ions) that can exist in the film, and form the diffuse component of the electric double layer, is limited by their "solubility" in hexyl alcohol. Rearranging equation (2) with $x = X$ we have

$$X = \frac{2}{\nu\sqrt{\beta c_0}} \cos^{-1} \sqrt{\frac{c_0}{c_X}} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

which reduces, when c_X/c_0 does not greatly exceed unity, to

$$X \doteqdot \Gamma/\nu\sqrt{c_X \, . \, c_0} \quad . \quad . \quad . \quad . \quad . \quad (7a)$$

The figures for X(calc.) in Table VI were obtained by taking $c_X = 5 \times 10^{-5}$ mole/*l*, and they agree very well with the observed thicknesses. If c_X were strictly constant and independent of p_0, as has been supposed, Γ must vary in accordance with the equation

$$\Gamma = 2\sqrt{\frac{c_X - c_0}{\beta}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

which may be obtained direct from equation (1), since when $x = X$ the potential gradient is $\frac{RT}{\nu F} \cdot \frac{1}{c} \frac{dc}{dx} = \frac{4\pi \Gamma F}{\epsilon}$. Accordingly at a value of p_0 for which $c_0 = c_X$, both Γ and X should vanish. This would occur in the case of hexyl alcohol on glass when $p_0 = 1{,}240$ dynes/cm.^{2} On the other hand, as p_0 and c_0 go to zero, Γ goes to the limit $2\sqrt{c_X/\beta}$, which in this case works out at $1{\cdot}94 \times 10^{-10}$ milliequivalent/cm.^{2} It will be noted that the valency ν of the counter ions does not enter into equation (8). This is because c is defined as the concentration in moles/*l*, not equivalents/*l*.

Considered as a whole, the results of Kussakov are in sufficient agreement with the theoretical predictions to indicate that the only important force acting between charged surfaces separated by distances of the order of 100 mμ is a repulsive force due to the osmotic activity of the ions in the intervening film. The most serious challenge to this conclusion arises from the results obtained with solutions of sodium chloride, which appear to require an additional repulsive force.

TABLE VI.

n-HEXYL ALCOHOL ON GLASS $\epsilon = 16$, $\nu = 1$.

p_0 (obs.), dynes/cm.^{2}	c_0 (calc.), moles/l.	$X_{\text{max.}}$ (calc.), mμ.	X(calc.), for $c_X = 5 \times 10^{-5}$ mμ.	Film Thickness (obs.), mμ.
257	$1{\cdot}04 \times 10^{-5}$	133	93	89
200	$0{\cdot}81 \times 10^{-5}$	151	111	108
167	$0{\cdot}67 \times 10^{-5}$	166	126	126
155	$0{\cdot}62 \times 10^{-5}$	173	134	135
150	$0{\cdot}59 \times 10^{-5}$	177	138	143

Summary.

Changes in the thickness of liquid films between solid surfaces play an important part in the swelling and shrinking of many systems.

The equation

$$X = \frac{2}{\nu\sqrt{\beta c_0}} \cot^{-1} \frac{2\sqrt{c_0}}{\Gamma\sqrt{\beta}}$$

has been obtained, in which X is either the thickness of a liquid film adhering to a single solid surface on which there are charged atom-groups at surface density Γ, or half the thickness of a film between two such surfaces. In obtaining the equation it is assumed that ions of one kind only, and of valency ν, constitute the diffuse component of the electric double layer, and that these ions are in ideal solution. β is a constant depending on the temperature and the dielectric constant. c_0 is the ionic concentration at the outer surface of the film on one surface or that in the centre of the film between two surfaces. It is considered that the excess hydrostatic pressure acting on the film when it is in equilibrium with the bulk liquid at the same height is RTc_0.

The calculated thicknesses of films of water on mica and glass, of hexyl alcohol on glass, and of hexane on steel are in good agreement with the optical measurements of Deryagin and Kussakov. A calculation on the same lines does not satisfactorily reproduce the experimental data for solutions of sodium chloride on glass.

Soil Physics Department,
Rothamsted Experimental Station,
Harpenden, Herts.

*GENERAL DISCUSSION**

Dr. J. Th. G. Overbeek (*Eindhoven*) said : A regular distance of 2000 A. between the plates of montmorillonite as mentioned by Dr. Mering may be explained by the repulsive action of the electrical double layers, if the swelling medium is pure water. If this explanation is valid the degree of swelling in solutions of electrolyte should be much smaller.

Dr. R. K. Schofield (*Harpenden*) said : So long as the height of the steps on the surface is small compared with the length of level areas, the liquid film should be able to follow the surface irregularity and give an average reflection substantially the same as for a truly flat surface.

Dr. D. M. C. MacEwan (*Harpenden*) (*communicated*) : I wish strongly to support Brindley's point about the especial value of clay minerals for studying the mechanism of swelling and shrinking. In connection with his remarks on halloysite hydration, I would like to point out that halloysite, and also the minerals of the montmorillonite group, show a type of adsorption reaction with organic liquids which is quite analogous to the swelling of organic high polymers (apparently, only *swelling* takes place —no phenomenon analogous to *solution* of polymers has been observed). The phenomenon is particularly easy to study with clay minerals, because the swollen materials all form definite complexes, giving X-ray diagrams with sharp lines, from the position of which the arrangement of the adsorbed organic molecules can be deduced. The good quality of the X-ray patterns is partly due to the fact that the clay minerals are polymerised in two directions to form large sheets of molecular thickness—a type of polymerisation which is rare among organic materials—and swelling takes place through adsorbed molecules going between the sheets.

An important point which has emerged from my recent work on these complexes (and the parallel work of Bradley [1]) is that the adsorbed molecules arrange themselves in parallel layers between the clay mineral sheets. These layers appear to be like two dimensional liquids, i.e. the molecules composing each layer have considerable mobility *within* the layer. This is suggested by the fact that the effective thickness of each layer always corresponds to the maximum width of the molecule in the orientation adopted, i.e. no packing takes place between layers. With halloysite, only one-layer complexes (similar to the hydrated form described by Brindley) have been obtained, and that only with molecules having many polar groups. With montmorillonite, complexes have been obtained with one, two, and three layers between the structural sheets. Fourier syntheses show that even in the three-layer complexes, the organisation in layers is extremely good. As the accompanying table shows, the complex-forming capacity of an organic molecule is roughly measured by the dipole moment divided by the molecular volume : the larger this quantity, the greater the number of layers in the complex. There are several exceptions to this general rule, however.

A word may be said here about the bearing of this work on Levine's, and Verwey and Overbeek's papers. The structural units of montmorillonite represent a very good practical realisation of the " charged plates " postulated by these authors. My results show that for a system of such parallel charged plates immersed in a suitable liquid, there must be a deep potential minimum at a separation varying from 10·4 A. (nitromethane) to 4·5 A. (1-propanol), so that the sharp monotonic fall in the potential curve for low separations, shown in Fig. 1 of Verwey and Overbeek's paper, cannot represent the true state of affairs. In fact, formulæ applicable to charged sheets surrounded by ionic clouds in a liquid medium probably cease to be valid for such close distances of approach, when the organisation of the inter-sheet molecules becomes quite different from that of an ordinary liquid; that is, when the adsorbed layers of the two charged sheets overlap. Levine's formula might lead to a potential minimum at about the inter-sheet distance found for some of the montmorillonite complexes, but for the same reason its applicability is doubtful. For very short distances of separation between the charged plates, the specific form of the liquid molecules (and probably also of the charged plates themselves) must clearly be taken into account in any explanation of the observed effects.

---

* On three preceding papers.

[1] Bradley, *J. Amer. Chem. Soc.*, 1945, 67, 975.

TABLE I.

Complex-forming Substance.	(μ/[P]) × 10^3.	No. of Layers in Complex with Montmorillonite.	No. of Layers in Complex with Halloysite.
Heptane	0	0	0
1-Heptanol	5·2	1	0
Ethanediamine	8·7	1	1
1-Propanol	9·9	1	0
1 : 3-Propanediol	12·0	2	1
Ethanol	12·8	2	1
Ethylene glycol	13·2	2	1
Nitrobenzene	15·9	2	0
Methanol	18·0	2	1
Acetone	18·0	2	1
Nitromethane	23·0	3	1
Acetonitrile	26·7	3	1

μ = Molecular dipole moment in Debye units (data mainly from Le Fèvre, *Dipole Moments*, Methuen, 1938).
[P] = Sugden's parachor, from data in Sidgwick, *The Electronic Theory of Valency*, O.U.P., 1927.

The potential minimum at larger distances is rather uncertain, both from theoretical and practical evidence: but if real, it follows that there may in some circumstances be two positions of equilibrium for a pair of parallel charged plates immersed in a liquid, one at a large distance (several hundred to several thousand A.) with ordinary liquid between them, and one at a very short distance (4-10 A.) with two-dimensional liquid layers between them. The experimental work mentioned here applies, of course, strictly speaking not to two, but to a series of parallel plates.

Dr. J. Th. G. Overbeek (*Eindhoven*) said: It may be pointed out, that the calculations of Dr. Schofield on a double layer with only one kind of ion present can be considered as a simple and useful approximation to the calculations on the more familiar type of double layer containing both negative and positive ions (Derjaguin, Langmuir, Verwey, Overbeek) for the case where the surface charge is large and the thickness of the layer is comparatively small.

The accord obtained in Tables I and II is indeed very satisfactory. The fact that the observed compressibility is somewhat larger than the calculated one might be ascribed to the influence of the double layer on the water-air interface. Another support for Dr. Schofield's theory is the fact mentioned by Derjaguin that the thickness of the layer on glass vanishes under the limit of observation by the addition of polyvalent ions like Al.

The bad accord in the case of monovalent ions (Table III) might seem alarming but I want to suggest that it is a peculiarity of glass perhaps connected with the swelling of its surface. The observed film thickness might be the thickness of the swollen layer rather than the thickness of the liquid film. Another indication for the swelling of glass surface is the observation of a very large surface conductivity, much larger than can be related to the properties of the double layer. It is desirable to repeat Derjaguin's experiments with electrolytes on a surface other than glass such as silver-halides, mica, etc.

Dr. G. W. Brindley (*Leeds*) said: In view of the very careful experiments carried out by Dr. Mering on the swelling of montmorillonite, and the present uncertainty regarding the indexing of the X-ray reflections from this mineral, it would be particularly valuable to have Dr. Mering's views on this point. The evidence at present points to a high degree of disorientation in the arrangement of successive layers and if this is correct then only basal (00*l*) reflections and two-dimensional (or cross-grating), (*hk*) reflections should occur. There should be no general reflections of type (*hkl*).

THE SWELLING OF PROTEIN FIBRES IN ORGANIC SOLVENTS.

By (the late) Dorothy Jordan Lloyd, Mary Dempsey and Marjorie Garrod.

Received 5th July, 1946.

A series of previous communications [1, 2, 3, 4] has dealt with the swelling of protein fibres in aqueous solutions. There is a large literature on this aspect of protein swelling, but the study of swelling in organic solvents has hitherto received but little attention.

The organic solvents examined in some detail in the present work were the fatty alcohols (first four members of the series), the fatty acids (again first four members), the substituted fatty acids thioglycollic acid and lactic acid, formamide and *m*-cresol: all these cause swelling. Other solvents studied in less detail were the higher members of the fatty alcohols and fatty acids: these have comparatively little effect; pyridine, dioxan, paraldehyde, benzene and ether: at room temperature these are without effect. Liquid ammonia (at the subliming temperature of solid carbon dioxide) is also without effect.

Protein fibres can conveniently be classified into two groups—those existing naturally in equilibrium with the atmosphere, which may be called "dry" fibres, and those existing naturally in equilibrium with the body fluids, which may be called "wet" fibres. The "dry" fibres examined in the present work were natural fibres of silk fibroin and keratin and artificial fibres of "Nylon" both after extrusion and after cold-drawing. The "wet" fibres were natural fibres of collagen, reticulin and elastin. Dry leaf gelatin was examined with the wet fibres.

The following material was used:

"*Nylon*" (for surgical sutures) both in the extruded and in the cold-drawn condition (kindly given by Imperial Chemical Industries (Plastics) Ltd.). The cold-drawn Nylon had been prepared from the extruded Nylon by stretching to about five-and-a-half times in length.

Silk fibroin.—Silk gut (for surgical sutures) prepared by cold-drawing the silk sac of the silk-worm.

Keratin.—Horsehair (for surgical sutures).

Collagen.—Tendons pulled from the tail of an adult rat and dried either by evaporation in air at room temperature, or by immersion in acetone with subsequent air-drying. These fibres are surrounded by a sheath of reticular fibres with some elastin fibres.

Elastin.—Prepared from *ligamentum nuchae* of the ox by the method of Schneider and Hajek.[5] Pieces of the ligament were heated under pressure in water at 120° for two hours to remove the collagen fibres (white fibres), boiled in several changes of distilled water until the washings no longer gave a biuret reaction, and finally dehydrated in acetone. The material was obtained in the form of a block of fibres lying roughly parallel. These fibres are surrounded

---

[1] Jordan Lloyd, Marriott and Pleass, *Trans. Faraday Soc.*, 1933, **29**, 554.

[2] Jordan Lloyd, Marriott and Pleass, *ibid.*, 1933, **29**, 1228.

[3] Jordan Lloyd, Marriott and Pleass, *ibid.*, 1934, **30**, 944.

[4] Jordan Lloyd, Marriott and Pleass, *ibid.*, 1936, **32**, 932.

[5] Schneider and Hajek, *Biochem. Z.*, 1928, **195**, 403.

by a fine sheath about which little is known except that it resists the action of trypsin longer than the interior of the fibre. Elastin fibres branch, and the branches pass into other fibres. Elastin from many parts of the body can be regarded as a network rather than as a collection of single fibres. In the case of *ligamentum nuchae* the "network" consists mainly of parallel fibres with a few connecting fibres. Besides the elastin fibres which were studied under the microscope, small thin slices of *ligamentum nuchae* (acetone dehydrated) cut along the main run of the fibres were also used for measurements of change of length (along the run of the fibres) and of width and thickness (across the run of the fibres), L, W and T in Table III. Considerable variation was found between W and T values, possibly partly due to the measurements being not truly across the fibres.

Reticulin.—Prepared from beef suet by extracting the suet in a Soxhlet apparatus with light petroleum until the membraneous part between the fat deposits was free from fat. The membrane was then dehydrated with acetone. The reticular tissue is made up of a network of fine fibres running in all directions. Individual fibres could not readily be teased out from the membrane, hence small rectangles were cut and measured in two directions at right angles ((*a*) and (*b*) in Table III) (which did not necessarily run along and across the fibres). It might have been anticipated, since the structure appears to be uniform, that dimensional changes would have been the same in the two directions. The differing results actually obtained may be due partly to unrecognised directional effects in the structure, or partly to uneven stretching when the membrane was smoothed out for drying after removal from the acetone.

Gelatin.—Coignet's gelatin cut into strips was used. From its *p*H swelling curve it appears to be an alkali-prepared gelatin.

The dry fibres were prepared for experiment by first drying at mild warmth (45°) and then storing for at least a week in a desiccator over fresh calcium chloride.

The "wet" fibres were used in the air-dry condition. This procedure is open to criticism, but thorough drying of proteins causes irreversible changes in their swelling properties. For instance, previous work[1] has shown that thorough drying by evaporation has a repressing effect on the swelling of collagen, possibly due to the formation of new cross-bonds between adjacent molecules. Drying gelatin gels by evaporation at room temperature considerably reduces the swelling capacity of the gels.[6] Further heating of the dry gelatin alters the character of gelatin. This is illustrated by the following :

	Solubility in		
	Warm water (60°)	*Formamide* (20°)	*Formic acid* (20°)
1. Untreated gelatin (control) . .	Dissolved slowly	Dissolved readily	Dissolved readily
2. Heated in air oven at 98° for 24 hours	Dissolved slowly	Dissolved slowly	Dissolved fairly readily
3. Heated in air oven at 140° for 24 hours	Swollen ; not dissolved	Swollen ; not dissolved	Partly dissolved

Liesegang[7] finds that heating gelatin above 100° destroys its power to be hydrolysed by trypsin, this again suggesting a structural change in the molecule due to heating with loss of moisture.

The organic solvents used in the experiment were freshly-opened bottles of the highest standard of purity obtainable. They were not dried further. Mardles[8] has found that the addition of water to organic liquids containing an OH or NH_2 group in the molecule will lead to their becoming solvents for gelatin, and quotes benzyl alcohol and pyridine. Speakman[9] also finds that water-soaking leads to an increase in the number of members in the alkyl series of organic acids that can penetrate a wool fibre. The amount of water considered by

---

[6] Jordan Lloyd, *Biochem. J.*, 1931, **25**, 1580.
[7] Liesegang, *Kolloid Z.*, 1942, **101**, 272.
[8] Mardles, *Biochem. J.*, 1924, **18**, 215.
[9] Speakman, *Proc. Roy. Soc. A*, 1931, **132**, 167.

Speakman and by Mardles is, however, considerably greater than the amount present as an impurity in any of the experimental liquids.

Swelling was followed by measuring length and diameter of the fibres before immersion and after various times of immersion. The diameter of the coarser fibres was measured in at least four places with a micrometer gauge, and the average taken. The diameter of the finer fibres was measured by a micrometer scale under the microscope, using a combination of lenses of known magnifying power. The larger fibres, e.g. tendons, were watched for an hour or so, then given two days in the organic liquid to reach equilibrium. The finer fibres, namely, elastin, were watched under the microscope. With the larger fibres, two or three fibres were taken in each experiment, and three to six measurements of diameter were made on each fibre before and after swelling. With elastin fibres, which on account of their small size could not readily be followed through the series as individuals, and which varied considerably in diameter, up to forty measurements of width were made to obtain an average. As mentioned above, reticulin was studied in the form of a membrane; the changes in dimensions of the individual reticulin fibres were not followed. Small pieces of reticular membrane and of *ligamentum nuchae* were measured for length and breadth, and the latter also for thickness.

The gelatin strips which, if they swell at all, swell in all three dimensions, were fixed edgeways under the microscope and the swelling measured as a change of thickness. Owing to the original conditions of drying, which tend to prevent shrinkage in area, the greatest shrinkage of the original gel is in thickness; hence the greatest swelling on the re-imbibition of a fluid is also in this dimension.

Swollen measurements were expressed as a percentage of initial dry measurements. The experiments were conducted within a range of temperature between 0° and 55°, the exact region chosen for the experiments varying for the different fibres as shown in the Tables. In studying temperature effect the same material was used throughout the experiment.

In considering the behaviour of any single class of fibre it is to be noted that there is individual variation within the class. This indicates the desirability of a statistical study. However, since the differences between classes were found to be considerably greater than the differences between individuals within the class, this was not undertaken. In view of this variability of the material, only first order effects were taken into consideration.

Collagen fibres (e.g. tendons) always exist surrounded by a reticular sheath. Radial swelling of a dry fibre, which leads only to about 100 per cent. increase in diameter, can occur without the rupture of this sheath, but radial swelling of the order of 200-1000 per cent. of the initial measurements always takes place with the partial rupture of the sheath, leaving regularly spaced rings which occur along the tendon.[10] Swelling of this order is always accompanied by marked axial contraction, so that the long thin fibre is transformed into a swollen glassy shortened structure, sometimes lobulated but sometimes not much longer than it is wide—in fact, the tendon appears to be under forces that are tending to bring it into a spherical form.

On account of this anatomical factor, which is difficult to control, the results of experiments with tendons always show individual variations. The following values show the final dimensions of collagen fibres in formic acid at 15° (expressed as a percentage of the initial length) with (1) a single collagen fibre with its reticular sheath and (2) a double collagen fibre with a larger amount of reticular tissue:

Fibre 1 (single)	length, as percentage of initial	22·2	22·5	15·2
	width ,, ,, ,, ...	2000	1760	2370
Fibre 2 (double)	length, ,, ,, ,, ...	29·7	27·5	27·5
	width, ,, ,, ,, ...	1300	1225	1000

Where the larger amount of reticular tissue occurs, radial swelling and axial contraction are both considerably inhibited.

Experimental Results.

These are shown in the Tables. Swelling had reached an apparent equilibrium after 24 hours, but in many cases 48 hours elapsed before measurements were taken, and in some cases a week or even longer. The figures shown in Tables I, II, III and IV are the results of continuous experiments, *i.e.* the same fibres were used over the whole temperature range for each solvent.

---

[10] Jordan Lloyd and Marriott, *Proc. Roy. Soc. B*, 1935, 118, 438.

TABLE I.—THE SWELLING OF "DRY" FIBRES IN WATER, ALCOHOLS AND FORMAMIDE AS PER CENT. OF INITIAL DRY MEASUREMENTS.

	Days in Solvent.	Temp.	Water.		Methyl Alcohol.		Ethyl Alcohol.		*n*-Propyl Alcohol.		*n*-Butyl Alcohol.		Formamide.	
			L.	W.	L.	W.	L.	W.	L.	W.	L.	W.	L.	W.
NYLON (UNSTRETCHED)														
Start ↑	1	25°	103	108	105	109	103	106	100	101	100	101	100	101
	1	30°	103	107	105	108	105	106	100	102	100	102	100	101
	1	37°	103	108	106	108	105	104	102	104	100	103	101	103
	1	45°	103	107	106	107	105	106	103	105	100	103	102	105
	4	25°	103	107	105	107	105	106	103	105	100	103	102	103
	1	55°	103	107	105	107	105	107	105	106	102	105	103	103
	5	37°	103	105	105	107	105	106	105	105	102	106	103	102
↓ Finish	113	25°	103	106	105	108	105	106	105	105	103	108	102	103
NYLON (COLD DRAWN)														
Start ↑		25°	102	104	103	106	103	107	100	102	100	100	100	100
		30°	102	108	103	106	103	109	100	103	100	100	100	100
		37°	102	108	103	106	103	106	101	105	100	100	102	100
		45°	102	107	103	108	103	104	102	104	100	100	102	102
		25°	102	104	103	105	103	108	102	105	100	100	102	100
		55°	102	104	103	108	103	106	103	105	101	102	102	102
		37°	102	103	103	107	103	107	103	103	101	103	102	104
↓ Finish		25°	101	104	103	107	102	104	102	102	101	105	102	100
SILK														
Start ↑		25°	100	115	100	117	101	100	100	101	100	100	100	106
		30°	101	112	100	115	101	101	100	100	100	100	100	140
		37°	101	114	100	117	101	101	100	102	100	102	101	139
		45°	101	115	100	116	101	100	100	102	100	103	101	133
		25°	101	114	100	117	101	104	100	102	100	100	101	130
		55°	101	117	100	116	101	106	100	100	100	100	101	131
		37°	101	112	100	116	101	111	100	102	100	100	101	130
↓ Finish		25°	100	113	100	116	—	—	100	102	100	101	100	147
			← flexible →				← stiff →						← flexible →	
HORSEHAIR														
Start ↑		25°	102	113	101	114	100	121	100	100	100	106	100	106
		30°	102	111	101	110	101	125	100	104	100	102	100	121
		37°	102	114	101	109	101	127	100	100	100	102	101	145
		45°	102	116	101	111	101	150	100	102	100	102	101	120
		25°	102	106	101	110	101	160	100	101	100	104	101	117
		55°	102	130	101	113	101	128	100	100	100	103	101	120
		37°	102	115	101	113	101	130	100	106	100	102	101	120
↓ Finish		25°	100	108	100	114	101	130	100	110	100	100	100	108
			← Slightly extensible and elastic →						← Not elastic →				← Elastic →	

L. = Length; W. = Width.

TABLE II.—THE SWELLING OF DRY FIBRES IN ORGANIC ACIDS OR META-CRESOL AS PER CENT. OF INITIAL DRY MEASUREMENTS.

	Days in Solvent.	Temp.	Formic Acid, $pK = 3{\cdot}7$.		Acetic Acid, $pK = 4{\cdot}7$.		Propionic Acid, $pK = 4{\cdot}8$.		Butyric Acid, $pK = 4{\cdot}8$.		Thioglycollic Acid, $pK = 3{\cdot}3$.		Lactic Acid, $pK = 3{\cdot}9$.		*Meta*-Cresol.	
			L.	W.	L.	W.	L.	W.	L.	W.	L	W.	L.	W.	L.	W.
Nylon (Unstretched) Start	1	25°	diss.		106	117	100	106	104	111	diss.		100	106	diss.	
	1	30°	diss.		106	117	102	112	108	114	diss.		103	113	diss.	
	1	37°	diss.		106	117	105	113	109	114	diss.		107	118	diss.	
	1	45°	diss.		106	117	106	113	109	114	diss.		110	120	diss.	
	4	25°	diss.		106	117	106	111	109	114	diss.		111	118	diss.	
	1	55°	diss.		106	117	106	113	109	114	diss.		112	120	diss.	
	5	37°	diss.		106	117	106	113	109	114	diss.		112	120	diss.	
Finish	113	25°	diss.		106	117	106	113	109	113	diss.		111	121	diss.	
Nylon (Cold Drawn) Start		25°	diss.		101	115	100	101	102	107	diss.		102	101	diss.	
		30°	diss.		102	118	100	104	103	108	diss.		103	103	diss.	
		37°	diss.		102	115	103	105	104	110	diss.		103	108	diss.	
		45°	diss.		102	115	104	105	104	109	diss.		103	110	diss.	
		25°	diss.		102	115	104	105	104	105	diss.		104	112	diss.	
		55°	diss.		102	119	104	109	104	109	diss.		103	114	diss.	
		37°	diss.		102	117	104	110	105	108	diss.		103	109	diss.	
Finish		25°	diss.		101	119	103	105	102	107	diss.		101	115	diss.	
Silk Start		25°	79	350	100	102	101	104	100	95	100	136	100	104	100	99
		30°	77	350	100	108	101	104	100	102	100	142	100	104	100	100
		37°	80	370	100	108	101	103	101	103	100	125	101	106	100	102
		45°	78	340	100	108	101	106	101	98	100	147	101	111	100	100
		25°	80	360	100	113	100	101	100	95	100	167	102	109	100	100
		55°	Mauve, very swollen, breaking up		100	127	101	102	100	103	100	150	103	130	100	104
		37°			101	124	101	105	101	100	100(soft)	134	103	130	100	101
Finish		25°			100	123	100	103	101	111	very soft breaking up		103	130	—	—
					← stiff →								flexible		stiff	
Horsehair Start		25°	100	156	100	117	100	102	101	113	99	150	100	106	100	114
		30°	100	145	100	140	101	106	100	130	98	150	100	116	100	110
		37°	100	160	100	140	101	106	100	130	97	170	100	130	100	110
		45°	100	190	100	140	101	126	100	130	103*	290	100	160	100	118
		25°	100	175	100	130	100	127	100	130	breaks	280	100	130	100	109
		55°	100	193	100	125	100	130	100	135	diss.		100	155	100	105
		37°	100	233	100	135	100	130	100	129	diss.		101	130	100	100
Finish		25°	Mauve, very stretchy and elastic. Splitting and tender		99	133	99	124	99	134	* very stretchy and elastic		elastic		Not elastic	
					← elastic →											

L = length; W = width; diss. = dissolves.

It will be noticed that with the fibres, in some cases there is a gain in length, in others no change in length, in others loss in length, while in *all* cases where the reagent penetrates there is a gain in diameter. Where there is loss of length the fibre develops rubber-like extensibility. With the small pieces of reticular membrane there may be a small gain in length and width which was interpreted as gain in the diameter of the individual fibres without appreciable change in length, or there may be loss of length and width (always accompanied by an obvious gain in thickness, though this was not measured) which was interpreted as loss in length with gain in diameter of the individual fibres. This interpretation was supported by the fact that such pieces showed rubber-like extensibility. With the pieces of *ligamentum nuchae* swelling was always in length, breadth and thickness, as in the gelatin pieces, and was accompanied by the development of rubber-like elasticity.

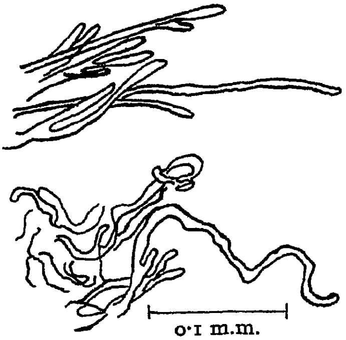


FIG. 1.—The coiling and swelling of elastin fibres (from ox *ligamentum nuchæ* in lactic acid). Upper drawing before treatment; lower drawing after treatment. Camera lucida drawing. In order to facilitate the drawing the lactic acid was run under the cover glass so that the movement of the fibres was actually somewhat restricted.

Swelling of the elastin fibres isolated from *ligamentum nuchae* was usually accompanied by coiling, as shown in Fig. 1 of the same fibres before and after treatment with lactic acid. The elastin fibres were not initially circular in cross-section. The shape is determined in the living animal by the close packing of the fibres in the ligament so that they become angular and three or more sided. The coiling during swelling is probably due to the release of strains caused by uneven compression of the individual fibres during growth. It was noted after swelling

TABLE III.—SWELLING OF "WET" FIBRES AS PERCENTAGE OF INITIAL DRY MEASUREMENTS.

20°.	Collagen Fibres.		Reticular Membrane.		Elastin Fibres.	Lig. Nuchæ (pieces).*			Gelatin.
	L.	W.	(a).	(b).	W.	L.	W.	T.	
Water . . .	100	200	112	110	116	113	140	110	400
Methyl alcohol .	100	140	102	104	120	112	108	109	114
Ethyl alcohol . .	100	120	—	—	118	107	111	117	105
n-Propyl alcohol .	100	110	—	—	100	105	109	100	—
n-Butyl alcohol .	100	100	—	—	100	106	100	120	—
iso-Butyl alcohol .	100	110	—	—	105	—	—	—	100
Formamide . .	25	550	110	105	134	128	142	137	diss.
Formic acid . .	diss.		30	70	160	150	240	230	diss.
Acetic acid . .	84	600	98	114	140	125	149	138	diss.
Thioglycollic acid .	diss.		85	89	140	150	270	170	diss.
Propionic acid .	100	150	100	112	134	116	129	130	260
Lactic acid . .	diss.		90	82	140	139	209	131	diss.
Butyric acid . .	100	290	100	105	126	123	149	123	130
Valeric acid . .	100	140	—	—	100	106	120	120	100
m-Cresol . .	—	—	—	—	—	132	201	170	diss.
40°									
Phenol . . .	28	410	100	108	—	150	150	—	diss.
80°									
Water . . .	24	550	80	90	—	110	130	—	diss.

L = length; W = width; T = thickness as percentage of initial dry measurements; diss. = dissolved.

* The pieces were measured at 15°, 20°, 30°, 45° and 10° in that order. The figures given are the average for the whole experiment in which the temperature effect appears to be limited to the rate of penetration of the solvent.

in formamide that although coiling occurred during swelling, the initial triangular cross-section of the fibre was retained, i.e. the fibre swelled unevenly in length but uniformly in diameter.

The results may be examined from five aspects :

(1) percentage change in length relative to percentage gain in width ;

(2) the effect of the lower homologues of a series (of solvents) compared with that of the higher homologues ;

(3) the effect of temperature ;

(4) the effect of the polarity of the solvent as indicated by its dielectric constant ,

(5) the conditions under which axial shrinkage occurs or swelling proceeds to solution.

Percentage Change in Length Relative to Percentage Change in Width. —Extruded " Nylon " fibres show swelling both in length and in diameter. This is to be associated with a highly non-orientated structure as in the gelatin strips. Elastin fibres also show swelling in all dimensions, and this may therefore reasonably be ascribed to the same cause. The three-dimensional swelling of elastin is shown in the experiments on the pieces of *ligamentum nuchae* (Table III) where the bulk of the fibres are lying in parallel alignment. Cold-drawn " Nylon " also shows slight swelling in length as well as diameter, though the ratio of percentage gain in length to percentage gain in width has been much reduced by the cold-drawing process, which considerably increases the degree of orientation in the structure, as shown by X-ray examination.[11]

TABLE IV.—SWELLING OF COLLAGEN FIBRES AS PERCENTAGE OF INITIAL DRY MEASUREMENTS.

Temp.	Water.		Formic Acid, $pK = 3{\cdot}7$.		Acetic Acid, $pK = 4{\cdot}7$.		Thioglycollic Acid, $pK = 3{\cdot}3$.	
	L.	W.	L.	W.	L.	W.	L.	W.
0°	—	—	16	3300	—	—	12	1500
15°	—	—	19	2900	104	430	12	1800
20°	100	163	18	4400	96	530	diss.	
25°	—	—	diss.		60	560		
30°	—	—			56	530		
37°	—	—			52	500		
45°	—	—			56	400		
55°	36	600			—	—		
80°					—	—		
Boiling	24	550			22	400		

Temp.	A Propionic Acid, $pK = 4{\cdot}8$.		Lactic Acid, $pK = 3{\cdot}9$.		Butyric Acid, $pK = 4{\cdot}8$.		Formamide.		*Meta*-Cresol.	
	L.	W.	L.	W.	L.	W.	L.	W.	L.	W.
0°	100	136	—	—	100	330	100	133	100	200
15°	100	240	35	1400	100	340	56	500	49	1000
20°	102	230	diss.		100	380	36	—	—	—
25°	100	195			100	370	38	530	41	1200
30°	102	210			100	380	36	500	36	2300
37°	100	170			100	340	34	715	34	500
45°	—	—			50	300				
55°	—	—								
80°										
Boiling	20	330			22	200	dissolves (ghost)		dissolves (ghost)	

L = Length ; W = width ; diss. = dissolves ; " ghost " = dissolves leaving the elastin and reticular sheath.

---

[11] Harris and Sookne, *J. Nat. Bur. Stand.*, 1941, **26**, 289.

Silk fibres do not show a gain in length greater than 1 % in any of the organic solvents, though the percentage gain in width may be as great as 50 (thioglycollic acid). Keratin fibres behave similarly. Collagen fibres show no gain in length in the alcohols, in propionic or butyric acid or in acetic acid at 15°, though the swollen width may be two or three times the initial dry width. Reticular fibres appear to behave similarly. This can be attributed to the high degree of orientation in the structure of the fibres. If the polymeric molecules are lying parallel to the fibre axis and fully extended, this would make further gain of length impossible, so that swelling, defined as an increase in volume, must occur as an increase in radial dimensions.

Under certain conditions, immersion in an organic solvent causes axial shrinkage in silk fibres, collagen fibres and reticular fibres. When this occurs it is accompanied by a high degree of radial swelling, which may lead to a three-and-a-half-fold (silk) or a five- or six-fold (collagen) increase in diameter. Where axial shrinkage occurs, rubber-like extensibility appears simultaneously.

The Effect of the Position of the Solvent in a Homologous Series.—The lower homologues obviously penetrate the fibres better than the higher homologues. Silk appears to be penetrated only by methyl alcohol and not by higher homologues, horsehair by methyl and ethyl alcohol but not by propyl and butyl, and cold-drawn Nylon by methyl, ethyl and propyl alcohol but not by butyl. Collagen, a natural "wet" fibre, after drying is penetrated by methyl, ethyl, propyl and *iso*-butyl alcohol, but not by butyl alcohol, and elastin after drying is penetrated by methyl and ethyl alcohol, but propyl alcohol appears to be the limiting case.

The failure of the solvent to penetrate may be attributed to two causes: (1) the molecule of the solvent becoming too large to penetrate the molecular grid of the fibre;[12] (2) diminishing polarity with increasing length of fatty chain. Polarity may be an important factor. The chemical nature of the solvent is also an important factor, since in many cases where the alcohol fails to penetrate the corresponding acid causes a high degree of swelling, e.g. extruded and cold-drawn Nylon in butyl alcohol and butyric acid respectively; also collagen and elastin in the same reagents.

On the "molecular grid" theory, the silk fibre would appear to have the lowest average pore size, since it is penetrated only by methyl alcohol among the alcohols. This suggests that in silk, unless the structure is affected by the solvent, the average pore size is 3-4 A. Speakman [9] has estimated the average pore size of the keratin fibre (wool) as about 6 A., which would agree with the observations on horsehair recorded here (i.e. penetration by methyl and ethyl alcohol with propyl alcohol forming a limiting case). Collagen and elastin fibres, which in their natural condition carry about 60 per cent. by weight of water, are readily penetrated by propyl alcohol, and may therefore have a more open structure, say 7-8 A. as the average pore size. However, there is no doubt that no solvent will penetrate any fibre unless it "wets" it, hence polarity must be an important factor.

The Effect of Temperature.—Temperature has little effect on the swelling in the lower homologues, but rising temperature favours swelling with the higher homologues. For instance, while the temperature was without effect on the swelling of extruded Nylon in methyl and ethyl alcohol, it had an effect in propyl and butyl alcohol, negligible swelling being found in these two solvents at 25° and 30°, and appreciable swelling at 37° and higher. After swelling at a higher temperature, lowering the temperature did not reduce the swelling. Again, horsehair showed swelling in formic and acetic acids at all temperatures, whereas in propionic acid it showed none at 25°, slight at 30° and 37° and considerable at 45° and higher. Again, after swelling at a higher temperature lowering the temperature did not reduce the swelling.

The temperature effect may partly be attributed to the effect of temperature on the viscosity of the solvent, as notably in the case of lactic acid, or to a reduction in the cohesive forces of the fibre, so that after a critical point the solvent is able to penetrate.

The Effect of the Polarity of the Solvent.—The organic solvents used in these experiments can be arranged in order of their dielectric constants as follows (values taken from International Critical Tables, temperature about 20°):

---

[12] Speakman, *Trans. Faraday Soc.*, 1930, 26, 61.

FORMAMIDE.	STRONG ORGANIC ACIDS.	ALCOHOLS.	WEAK ORGANIC ACIDS.	*m*-CRESOL.
> 84	(*p*K 3·3-3·9)		(*p*K 4·7-4·8)	5·0
	formic acid	methyl	acetic	
	47·9	35	6·5	
	lactic acid	ethyl	propionic	
	19-23	27	3·1	
	thioglycollic acid	propyl	butyric	
	?	22	2·8	
		iso-butyl	valeric	
		19	2·6	
		n-butyl		
		18		

Though polarity appears to be related to swelling within a homologous group of solvents, there is no simple relation between polarity and swelling in passing from one group to another. Moreover, the relation between the dielectric constant and the solvent effect is not a simple one: gelatin is soluble at 20° in formamide, *m*-cresol, the strong organic acids and acetic acid: extruded "Nylon" in formic acid and *m*-cresol; collagen in formic, thioglycollic and lactic acid, but not in any of the alcohols.

Formic acid, however, with its high dielectric constant, has a greater effect on swelling than thioglycollic acid, even though the latter has a lower *p*K value, and both these acids have a greater effect on all the fibres than any of the other acids used in the experiments.

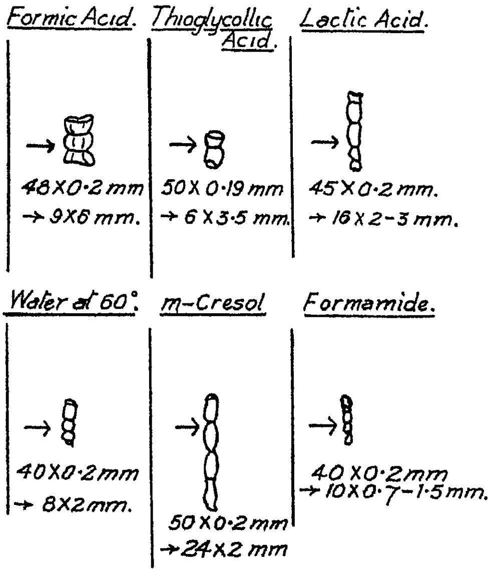


FIG. 2.—Swelling of collagen. Organic reagents at 20°. Water at 60°.

The figures for swelling in the polar solvent, water, have been included in the tables. The swelling in water does not differ noticeably from the swelling in methyl alcohol in the case of the "dry fibres", and of elastin from among the "wet" fibres. Collagen, reticulin and gelatin, all "wet" fibres, are alike in showing very much greater swelling in water than in methyl alcohol.

In formamide, the most polar solvent, the two forms of "Nylon" and the reticulin show little swelling. The other protein fibres show marked swelling, frequently greater than in water (silk, collagen), but not always greater than in the slightly polar solvent *m*-cresol (hair, collagen).

The Conditions Under Which Axial Contraction Occurs or Under Which Swelling Proceeds to Solution. —It is a familiar fact that collagen fibres in water undergo sudden and rapid loss of length when the temperature of the water exceeds a certain critical value. The temperature at which this happens depends on the *p*H of the system,[13] but as it generally occurs above blood heat it has come to be called "thermal shrinkage." Silk, collagen and reticulin fibres all, under certain conditions, show this change in organic solvents. The only condition under which the silk fibre showed axial contraction was in pure, freshly opened glacial formic acid, the change occurring over the range 25-45° (lower temperatures were not examined). In this temperature range the swollen and contracted fibre was stable; at 55° it became excessively swollen, tender, mauve in colour and broke up into fragments on touching. Collagen fibres showed axial contraction in formic, thioglycollic and lactic acids at 0° and 15°, and dissolved in these solvents at higher temperatures; they showed contraction in acetic at 20-25° and butyric at 50°, and further contraction at the boiling point, but did not dissolve; they showed axial contraction in formamide and *meta*-cresol at 15° and higher temperatures, and dissolved on boiling. The appearance of collagen fibres immersed in hot water or in various organic solvents at room temperature is shown in Fig. 2. The solvents which brought about axial contraction in collagen at room temperature all dissolved gelatin at the same temperature. These solvents also appeared to cause axial contraction in reticular fibres, but the latter did not dissolve under conditions which brought about solution of the collagen.

It seems clear, however, that the axially contracted state can be looked on as an intermediate condition between the natural state of the protein fibre and the state in solution. As mentioned earlier, it is characterised by a rubber-like extensibility.

Experimental evidence given here shows that the change non-extensible form ⇌ rubber-like form is reversible, but is accompanied by a deformation of the structure which is not entirely reversible. Clark and Schaad [14] have recorded that the thermal and chemical shortening of collagen completely abolishes the long meridional spacing shown by X-rays, and this is not restored by extending the shortened fibre.

Experiments 1, 2 and 3 described below deal with the silk fibre.

EXPERIMENT 1. SILK FIBRE IN FORMIC ACID → ALCOHOL: FIBRE NOT LOADED.*

		Long-range Elasticity.	Per Cent. of Initial Dry Measurements.		
			Length.	Width.	Volume (calc.).
Start ↑	Dry silk fibre I	absent (white)	100	100	100
	After 48 hours in glacial formic at 0°	present (mauve)	70	480	1100
	After 48 hours in first ethyl alcohol bath at 20°	absent (yellowish)	70	250	430
	After 3 days in second ethyl alcohol bath at 20°	absent (yellowish)	72	210	320
	After 2 months in third ethyl alcohol bath at 20°	absent (yellowish)	71	200	280
↓ Finish	Dried in air-oven at 45° and desiccator	absent (yellowish)	70	120	105

* The original dry fibre was 40 mm. long, 0·49 mm. diameter; volume = 7·55 c.mm. Volume of experimental bath = 20 cc. The final dry fibre was 28 mm. long. 0·60 mm. diameter; volume = 7·92 c.mm.

---

[13] Theis and Esterley, *J. Amer. Leather Chem. Assoc.*, 1940, **35**, 563.

[14] Clark and Schaad, *Radiology*, 1936, **27**, 339.

The axial contraction that is brought about in silk is therefore not reversed simply by washing out the formic acid, though this process removes the long-range elasticity. Hence new lateral links can form while the fibre is in the contracted condition. The axial contraction can be overcome by loading the fibre; loading by itself does not remove the long-range elasticity, but washing out the loaded fibre with ethyl alcohol does so. New lateral links, therefore, can reform while the fibre is in the extended condition.

EXPERIMENT 2. SILK FIBRE IN FORMIC ACID → ALCOHOL: *
FIBRE LOADED TO RESTORE EXTENSION.

		Long-range Elasticity.	Per Cent. of Initial Dry Measurements.		
			Length.	Width.	Volume (calc.).
Start	Dry silk fibre 2	Absent	100	100	100
↑	After 3 hours in glacial formic at 15°	present	75	340	860
	After loading to original length with 54 g. weight	present	100	—	—
↓	After 3 hours in ethyl alcohol	absent	100	160	250
Finish	After drying	absent	100	102	104

* The original dry fibre was 40 mm. long, 0·46 mm. diameter; volume = 6·68 c.mm. Volume of experimental bath = 300 cc. The final dry fibre was 40 mm. long, 0·47 mm. diameter; volume = 6·95 c.mm.

It is interesting to notice that in both cases the dry fibroin fibre has the same volume at the end of the experiment as at the beginning, although with the unloaded fibre there has been a permanent loss of length with gain of width. The structure, however, has certainly undergone some permanent deformation since ethyl alcohol does not penetrate a dry silk fibre and cause it to swell. Similar evidence of permanent change in the fibre structure is shown by passing a silk fibre from formic acid to water. The volume in water after this treatment is considerably greater than if the dry fibre is passed into water directly, although

EXPERIMENT 3. SILK FIBRE IN FORMIC ACID → WATER. FIBRE NOT LOADED.

	Length.	Width (mm.).	Volume.	Per Cent. of Initial Measurements.		
				Length.	Width.	Volume (calc).
Dry silk fibre 3 . .	55	0·039	4·10	100	100	100
Fibre 3 after 24 hours in water at 15° .	55	0·374	6·04	100	121	147
Dry silk fibre 4 . .	50·5	0·408	6·60	100	100	100
Fibre 4 after 24 hours in water at 15° .	50·5	0·465	8·58	100	114	130
Dry silk fibre 5 . .	35	0·50	6·87	100	100	100
Fibre 5 after 28 hours in formic acid at 15° (elastic) . . .	27·5	1·80	70·0	78	360	1000
Fibre 5 after 24 hours in water at 15° (very slightly elastic) .	28·0	1·07	25·2	80	214	367

water penetrates the fibre readily. Note that water after formic acid largely, but not entirely, removes the long-range elasticity.

The long-range elasticity conferred on the collagen fibre by formic acid (and other reagents) can also be removed by washing out the formic acid, though the disturbance in the molecular structure caused by the change to the rubber-like form appears to be even greater than in silk, since at the end of the experiment the dry fibre has not returned to its original volume.

EXPERIMENT 4. COLLAGEN FIBRE IN FORMIC ACID → ALCOHOL: FIBRE NOT LOADED.*

		Long-range Elasticity.	Per Cent. of Initial Measurements.		
			Length.	Width.	Volume (calc.).
Start	Dry collagen fibre 1	absent	100	100	100
↑	After 48 hours in formic acid at 0°	present	16	3300	18,000
	After 48 hours in first ethyl alcohol bath at 20° . .	absent	6	1600	1,900
	After 48 hours in second ethyl alcohol bath at 20° .	absent	4	1200	540
↓	After 48 hours in third ethyl alcohol bath at 20° .	absent	4	1130	510
Finish	After drying in air-oven at 45° and in desiccator . .	absent	4	780	230

* The original dry fibre was 50 mm. long, 0·18 mm. wide; after formic acid bath was 8 mm. long and 6 mm. wide; after third alcohol bath was 2·05 mm. long and 2·03 mm. wide; after drying was 1·9 mm. long and 1·4 mm. wide. Volume of experimental bath = 20 cc.

EXPERIMENT 5. COLLAGEN IN FORMIC ACID → ALCOHOL: FIBRE LOADED.

	Length.	Width (mm.).	Volume.	Per Cent. of Initial Dry Measurements.		
				Length.	Width.	Volume (calc.).
Dry collagen fibre 2 .	41	0·14	0·63	100	100	100
Fibre after 3 hours in formic acid at 15° (elastic) . .	11	1·5	19·43	27	1070	3100
Fibre loaded to full extension (54 g.) .	42	—	—	103	—	—
Fibre after 3 hours in ethyl alcohol .	42	0·19	1·19	103	136	189
After drying . .	42	0·14	0·65	103	100	103

In experiment 4 there is actual further loss of length on replacing the formic acid with alcohol, though a dry collagen fibre placed direct into ethyl alcohol swells without change of length. If the formic acid-swollen fibre is loaded so as to restore its original length, subsequent washing out with alcohol will remove the long-range elasticity without change of the loaded length. After drying, the fibre will return to its original dimensions.

Under some circumstances the contracted rubber-like collagen fibre may show some relaxation without being submitted to tension. This occurs when a fibre axially contracted in hot water is put into cold water. The relaxation, however, is only partial, the fibre having returned to 45 % of its original dry length after four days, and the rubber-like character is retained. There appears also to be some relaxation when a fibre contracted in formamide is passed into cold water, the fibre returning to 62 % of its initial length after four days. Again the cold water fails to remove the rubber-like condition induced by immersion in formamide. These results are shown in the experiments recorded below, which also show that formamide after formic acid has an effect which is additive rather than substitutional for the first minute after immersion in formamide, but after this the fibre relaxes to 37 per cent. of its original length. Formic acid after formamide does not appear to cause any further contraction, but the fibre relaxes to 40 % of its original length before disintegrating.

The axially contracted state has not been found in "Nylon" or elastin fibres, but has been described as the "super-contracted" form of keratin.[15, 16] It is not shown by horsehair in any of the organic solvents used in this work, but horsehair dissolves in thioglycollic acid at temperatures above 45°, doubtless because this reagent reduces the disulphide bond as well as opening up the salt link.

EXPERIMENT 6. COLLAGEN FIBRES: NOT LOADED.

	Length (mm.).	Per Cent. Initial Dry Length.
HOT WATER → COLD WATER		
Dry collagen fibre 3. Initial	90·0	—
After water at 80° c. (elastic)	22·0	24·5
In water at 20° c. (elastic)—		
After 1 minute	26·0	28·8
„ 5 „	31·0	34·4
„ 10 „	33·0	36·7
„ 30 „	34·0	37·7
„ 1 hour	36·0	40·0
„ 3 „	38·0	42·2
„ 5 „	38·0	42·2
„ 24 „	39·0	43·3
„ 3 days	40·5	45·0
„ 4 „	41·0	45·5
„ 5 „	41·0	45·5
„ 6 „	41·0	45·5
FORMAMIDE → COLD WATER		
Dry collagen fibre 4. Initial	60·0	—
After formamide at 20° c. (elastic)	20·5	34·3
In water at 20° c. (elastic)—		
After 1 minute	22·0	36·6
„ 5 „	29·0	48·4
„ 15 „	31·0	51·7
„ 1 hour	35·5	59·1
„ 2 „	36·0	60·0
„ 3 „	36·0	60·0
„ 24 „	36·5	60·8
„ 3 days	37·0	61·6
„ 4 „	37·0	61·6
„ 5 „	37·0	61·6
„ 6 „	37·0	61·6

---

[15] Speakman, *Trans. Faraday Soc.* 1929, **25**, 169; *Nature*, 1929, **124**, 948.
[16] Astbury and Woods, *Phil. Trans. Roy. Soc. A*, 1933, **232**, 333.

EXPERIMENT 6. COLLAGEN FIBRES: NOT LOADED.—*Continued.*

	Length (mm.).	Per Cent. Initial Dry Length.
FORMAMIDE → FORMIC ACID		
Dry collagen fibre 5. Initial	60·0	—
After formamide at 20° c. (elastic)	21·5	35·8
In formic acid at 20° c. (elastic, glassy)—		
After 1 minute	21·5	35·8
,, 5 ,,	22·0	36·6
,, 15 ,,	22·5	37·5
,, 1 hour	23·0	38·3
,, 2 ,,	23·0	38·3
,, 3 ,,	23·5	39·2
,, 24 ,,	24·0	40·0
,, 3 days	Disintegrated	—
FORMIC ACID → FORMAMIDE		
Dry collagen fibre 6. Initial	70·0	—
After formic acid at 20° c. (elastic)	21·0	30·0
In formamide at 20° c (elastic, glassy)—		
After 1 minute	20·0	28·6
,, 5 ,,	20·5	29·3
,, 10 ,,	21·0	30·0
,, 15 ,,	22·0	31·4
,, 30 ,,	22·0	31·4
,, 60 ,,	23·0	32·8
,, 3 hours	24·5	35·0
,, 5 ,,	25·0	35·7
,, 24 ,,	25·0	35·7
,, 3 days	26·0	37·2
,, 4 ,,	26·0	37·2
,, 5 ,,	26·0	37·2
,, 6 ,,	26·0	37·2

Theory of Swelling.

The swelling of protein gels in aqueous systems of acids and alkalies has been satisfactorily explained on a simple theory, at least as far as volume and weight changes are concerned. Proteins are ampholytes, and at *p*H values removed from the iso-electric point form colloidal ions, hence in protein gels and fibres a membrane equilibrium will be established and swelling will occur until the internal excess osmotic pressure drawing water into the system is balanced by the cohesive forces. The effect of salts in repressing acid swelling is due to the suppression of the internal excess osmotic pressure. The effect of the salts of the lyotropic series in promoting swelling in solutions of nearly neutral reaction has been ascribed to a weakening of these cohesive forces.

This theory does not offer any explanation of the thermal shrinkage of collagen fibres in aqueous systems, which is accompanied by a gain in diameter and volume, and it obviously does not apply to the swelling that takes place in organic solvents. While there may be a membrane equilibrium set up in formic acid, since this acid forms highly ionised protein salts, there will not be in acetic or lactic acid, since the protein salts of these acids are not highly ionised.[17]

A comparison of the swelling of the different fibres in the different solvents shows that not only the structural pattern of the fibre but also the chemical constitution of the fibre and the chemical nature of the solvent must both be factors of importance (see Fig. 3, showing the different behaviour of various fibres under the same conditions). The cohesive forces of the fibre are due to lateral bonding between polymeric and generally orientated molecules, and the nature of the lateral bonds is characteristic for each protein, the disulphide link and the salt link being, for instance, typical of the keratin fibre. The effect of the solvents on the lateral bonds in a fibre will obviously, therefore, be a major factor in the occurrence of swelling. Thus the effect of a solvent such as formic acid acting on the electrovalent salt-link will result in a rupture of the link, with a subsequent crumpling of those portions of the fibre backbones held together by this type of link. The overall result will be an axial contraction to a position of greater entropy (the more stable condition), with a corresponding swelling as the adjacent backbones may be forced further apart by rupture of some of the lateral links.

---

[17] Greenberg and Larson, *J. Physic. Chem.*, 1935, **39**, 665.

A solvent with a high dielectric constant, such as formamide, should theoretically exert a large disrupting action on the electrostatic forces of the salt link, which may not be broken in the chemical sense, but the forces of attraction may be so weakened that the molecular backbones may be forced further apart, resulting in a swelling of the fibre.

The experiments described in this report have led to the following general theory: in " Nylon " and the protein fibres the long polymeric molecules may lie: (1) mainly in a crumpled or random configuration, in which case, provided enough solvent is present in the system to act as a lubricant, the fibre will be extensible if a stretching force is applied, e.g., extruded Nylon and elastin. If " Nylon " is stretched, the alignment of the molecules allows the formation of lateral bonds (direct carbimino links between the molecular backbones), so that on the release of the stretching force there is no reversion to the original length; if elastin is stretched, in spite of the alignment of the molecules there is no formation of lateral bonds (since there are few polar side-chains and the non-polar are mostly long, thus holding the backbones apart), so that on the release of the stretching force there is reversion to the original length. In these fibres, in which there is little molecular orientation, swelling is in three dimensions; (2) mainly in the extended condition, in which case, even if a solvent is present in the system, the fibre will be considerably less extensible if a stretching force is applied, e.g. cold-drawn " Nylon," silk, collagen and reticulin. Since the kinetic energy of the long polymeric molecule will not allow it, if free, to lie, on average, in the fully extended position, the stability of these fibres must be brought about by lateral bonding (either direct backbone linkage or through side-chains). If all the lateral bonds are of one type, conditions which open the bonds will give freedom to the molecules building up the fibre, which will dissolve

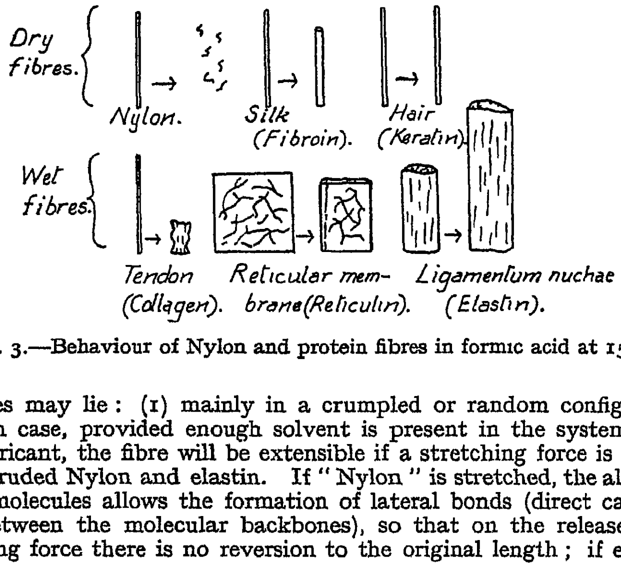


FIG. 3.—Behaviour of Nylon and protein fibres in formic acid at 15° C.

(cold-drawn "Nylon" in *m*-cresol). If the lateral bonds are of more than one type, conditions which open one type of bond only will free only sections of the long polymeric molecule; these will take on a reduced overall length with gain in lateral dimensions and entropy, and the development of rubber-like extensibility (silk in formic acid; collagen and reticulin in various solvents). The stability of the fibre will be maintained by the other type or types of lateral bonds, but when all types of bonds are opened the fibre will dissolve (silk in formic acid at 55°, collagen in formic acid at 20°). In fibres with molecules orientated parallel to the fibre axis, swelling will be radial only. (3) Mainly in an ordered but looped configuration, which again can only be maintained by lateral bonds. Here the opening of one type of bond will not necessarily lead to a shortening of the fibre, but will bring about rubber-like extensibility (keratin fibres in water or formamide). Axial contraction, however, can occur in such fibres, *e.g.* keratin fibres in steam [15] or in concentrated aqueous solutions of strong acids and alkalies.[3]

The chemical relation between the active groups in the different solvents and those in the different fibres will be dealt with in a separate communication.

Summary.

1. Protein fibres (and "Nylon") may swell in organic solvents. The degree of swelling depends on the nature of the fibre, the chemical class of the solvent, and, within a class, of the position of the solvent in a homologous series; the lower the homologue the greater the swelling.

2. Extruded "Nylon" and elastin resemble gelatin in that swelling occurs both in length and diameter. This is to be interpreted as due to disorientated molecular structure. Isolated elastic fibres coil when they swell, probably due to their uneven packing during growth. Cold-drawn "Nylon" silk, horsehair, collagen and reticulin fibres swell only (or mainly) radially. This is to be interpreted as due to orientated molecular structure.

3. Under certain conditions silk, collagen and reticulin fibres show marked axial contraction accompanying marked radial swelling. The fibres in this condition show rubber-like extensibility. The axially contracted state is interpreted as a half-way (though stable) stage towards solution. "Nylon" fibres never show this condition, but pass directly into solution.

4. Experimental evidence is given which shows that the change from the non-extensible form to the rubber-like form is reversible, but is accompanied by a deformation of structure which is not reversible. If the reagent which produced contraction is washed out with alcohol, new lateral bonds can reform while the fibre is in the fully contracted or in the loaded condition.

5. Temperature influences swelling in that many solvents appear to have a critical temperature below which they cannot enter the fibre.

6. Neither the size nor the polarity of the solvent molecules is sufficient to explain the interaction between solvent and fibre, and an explanation taking into account chemical interactions between solvent and localised groups in the fibre must be sought.

7. It is suggested that in fibres the polymeric molecules may be (1) disorientated as in extruded "Nylon," elastin, and as in gelatin; (2) orientated in parallel alignment and held by cross-bonds, as in cold-drawn "Nylon," silk, collagen and reticulin; (3) orientated in parallel but regularly folded alignment, as in keratin, and again held by cross-bonds. When all the cross-bonds are of one type, as in extruded "Nylon," a solvent opening this bond will dissolve the fibre; when the cross-bonds are of two or more types, a solvent which opens one type but leaves the others intact will (at least in fibres where the intact cross-bonds are widely scattered) lead to a freeing of sections of the long backbone, which will crumple with gain

of entropy, loss of overall length, and development of rubber-like extensibility; while a solvent opening all types of link will lead to a solution of the fibre. This type of swelling with axial contraction occurs in silk, collagen and reticulin.

The authors are indebted to the Department of Scientific and Industrial Research and to the Chairman and Council of the British Leather Manufacturers' Research Association for permission to publish this work.

RELATION OF SWELLING TO SHRINKAGE TEMPERATURE OF COLLAGEN.

By Edwin R. Theis.

Received 5th July, 1946.

Fahrion [1] in 1908 was the first investigator to discuss the behaviour of animal skin in boiling water as a quantitative measure of "tanning." He stated that under such treatment, ordinary hide powder is converted into glue, chamois leather is but little affected and chrome tanned leather is affected not at all. Ewald [2] later measured both the shrinkage temperature and the shrinkage percentage of connective tissue taken from mouse tails and frog leg sinews. His investigation indicated that acid treatment decreased the shrinkage temperature while the tanning procedure increased this factor. In 1924 Powarin and Aggeew [3] utilised this specific test in determining the actual temperature at which leather began to undergo thermal change. These investigators found that the normal "liming" procedure changed the thermal lability from 66·7° c. to 42·2° c. Chater [4] made a series of studies from which he concluded the following: (*a*) alkaline treatment reduced the normal shrinkage temperature; (*b*) preliminary soaking in acid and alkali solutions gave shrinkage temperatures first increasing toward the neutral zone and then decreasing in the more alkaline region; (*c*) he correlated his data with swelling measurements and noted specific trends; and (*d*) he concluded that the shrinkage temperature would be an excellent index for detecting any departure from a usual tanning procedure. Chambard and Michallet,[5] Schiaparelli,[6] Casaburi and Corradini [7] made use of the shrinkage temperature in investigations of various types of tannage. Theis *et al.*[8] used this factor in their studies of formaldehyde, vegetable and chrome tannages. Braybrooks *et al.*[9] in 1939 made an extensive study of the relation of shrinkage temperature to pickling and tanning and suggested important explanations of their findings. This work will be discussed further later in this paper.

In collagen, two types of cohesive forces must obtain: electrovalent or salt-like linkage and the co-ordinate or hydrogen bond. Thus in collagen the bonding together of the long polypeptide chains should be largely the interaction between the numerous peptide linkages and to a very small extent between the charged centres. Collagen chains are not fully extended as is the case of silk fibroin but are believed to be slightly curled. The presence of bulky R— groups, such as arginine, lysine, proline or oxyproline in *cis* and *trans* configuration may be the answer for the slight contraction of the collagen polypeptide chains. The studies of Astbury showed 2·8° A. per amino acid residue for the collagens in comparison to 3·5° A. per amino acid residue in silk fibroins. In collagen, the packing of the individual units is not so close as in the case of silk fibroin and wool keratin. Since natural or native collagen is held in compact form by co-ordinate and electrovalent forces, thermal agitation weakens these forces and the protein tends to revert to such a form not requiring strict parallel alignment of the polypeptide chains. Astbury [10] has stated " when the thermal agitation is sufficient to overcome the interchain attractions, the chain bundles may be said to melt and the chains collapse upon themselves. It follows that anything that interferes with the solution and interactions of the side chains that form the ring of the polypeptide grid, or with any inter-chain linkage for that matter, must inevitably influence the thermal transformation temperature."

---

[1] Fahrion, *Collegeum*, 1908, **338-39**, 459.

[2] Ewald, *Physiol. Chem.*, 1919, **105**, 115.

[3] Powarin and Aggeew, *Collegeum*, 1924, **650**, 101.

[4] Chater, *J. Int. Soc. Leather Trades Chem.*, 1929, **13**, 24; 1929, **13**, 427; 1930, **14**, 28; 1930, **14**, 133.

[5] Chambard and Michallet, *Le Cuir Tech.*, 1927, **16**, 520.

[6] Schiaparelli and Carxeggio, *ibid.*, 1925, **14**, 68.

[7] Casaburi and Corradini, *ibid.*, 1937, **26**, 20.

[8] Theis and Schaffer, *J. Amer. Leather Chem. Assoc.*, 1936, **31**, 515; Theis and Kalb, *ibid.*, 1938, **33**, 120; Theis and Steinhardt, *ibid.*, 1942, **37**, 433.

[9] Braybrooks, McCandlish, and Atkin, *J. Int. Soc. Leather Trades Chem.*, 1939, **23**, 111, 135.

Gustavson [11] states that in the shrinkage of collagen, in the presence of water, a disorganisation of the protein structure obtains. Kuntzel [12] has stated this type of protein shrinkage to be a melting-point mechanism —an irreversible change of collagen I to collagen II. Kuntzel has named this change " *verleimung*." Grassmann [13] claims that the inertness of native collagen toward the action of trypsin is destroyed by thermal treatment and thus rupture of stabilising effective linkages is indicated.

It has been stated that the binding together of the collagen polypeptide chains is due to co-ordinate and electrovalent forces. Jordan Lloyd [14] has extended the hypothesis of Astbury and Speakman, relative to the keratins, to collagen, namely, the presence of salt-like linkages as intermolecular forces formed through mutual attraction of opposite charged side-chains in juxtaposition on adjacent peptide chains (—COO^- ^+H_3N—). Gustavson [11] points out that in view of the given frequencies for the —COO^- group and further as the distribution of the interacting amino and carboxyl groups must conform to the most favourable steric conditions for the effective formation of salt-type linkage, in view of the expected sparse interspacing of such weak ionic linkages, these could not be the main cause of the strong intermolecular cohesive forces of collagen. While Kuntzel, Jordan Lloyd, Astbury, Speakman, Braybrooks, Theis and others recognise both the ionic type of cross-linkages and the intermolecularly co-ordinated peptide (hydrogen) bonds as the main connecting forces, Gustavson states that the mere discharge of the —COO^- group compensated by the ionic basic group does not provide a satisfactory explanation of acid swelling and shrinkage of collagen.

Braybrooks, McCandlish and Atkin [9] have shown that shrinkage temperature and swelling are a function of the hydrogen ion concentration of the medium and are parallel in trend. These investigators interpret their findings as a rupture of salt-like linkages by H^+ and OH^- ions.

Shrinkage temperature may be defined as the specific point at which the increasing disruptive tendencies exceed the diminishing cohesive forces, thus making the shrinkage temperature actually a measure of the structural stability of the collagen expressed in temperature units.

The term " hydrogen bond " or " bridge " will be frequently used but in all fairness it should be stated that " hydrogen bond " is still somewhat a matter of speculation in protein structure. Astbury [10] has particularly stressed this point. "As things stand at the moment, the question of hydrogen bond in protein structure is not at all one on which to embark with any confidence." When the word " hydrogen bond " is used in the work to follow, advantage has been taken of current knowledge to bring about clarity of explanation of certain research studies.

---

[10] Astbury, *J. Inter. Leather Trades Chem.*, 1940, **24**, 69.
[11] Gustavson, *J. Amer. Leather Chem. Assoc.*, 1946, **41**, 47.
[12] Kuntzel, *Strasny Festschript* (Darmstadt, 1937), 191.
[13] Grassmann, *Kolloid Z.*, 1936, **77**, 205.
[14] Lloyd and Shore, *Chemistry of Proteins*, 2nd Ed. (London, 1939).

From 1936 to date, Theis *et al.* [8] have been investigating the shrinkage temperature of collagen under many different conditions. The data shown graphically in Fig. 1 illustrates the effect of acid and base upon the swelling and shrinkage temperature of native collagen. The preparation of the collagen for such studies has been described elsewhere.[8] This figure indicates the structural stability as measured by the shrinkage temperature over the *p*H range 1·0 to 13·5 with and without salt addition. In addition Fig. 1 indicates the relation of swelling to structural stability. Curve A, representing acid or base solutions without added salt, indicates that mild acid or alkali (*p*H 5·0-9·0) has little or no effect upon the structural stability (hydrothermal stability) of the protein but that acid or alkali solutions of *p*H value less than 4·0 or greater than 9·0 have a drastic effect causing a definite decrease in shrinkage temperature and thus indicative of structural breakdown in these particular zones.

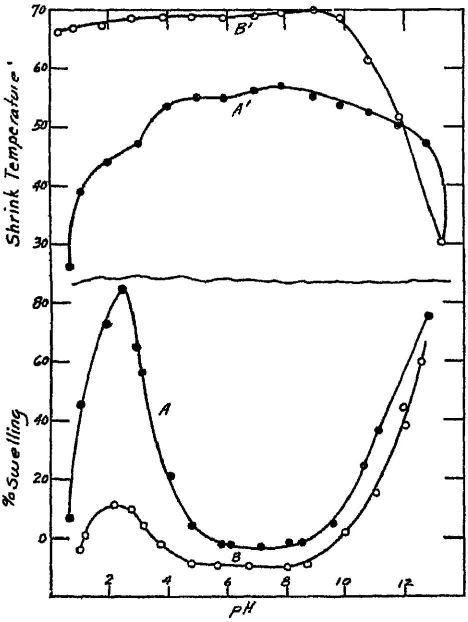


FIG. 1.

These facts make it rather evident that the action, of the acid in the low *p*H range and the alkali in the high range, is upon the short link or hydrogen bridge. Careful examination of the curve indicates that in the isoelectric or buffer zone, back titration of histidine obtains and since in collagen the amount of histidine is small, the effect is not great. At *p*H values greater than 10·0, lysine and arginine are affected and drastic decrease in hydrothermal stability result. On the acid side, aspartic and glutamic acids are affected, being back titrated and probably reaching approximate maximum values at *p*H 2·0. Further data indicate that at *p*H 3·0 approximately 70 per cent. of the acid-fixing group have been satisfied—indicative of a reasonable destruction of salt-like linkages. The same reasoning will obtain at *p*H values greater than 11·0. It is entirely logical to assume in the *p*H range 3·0 to 11·0 that most of the changes in structural stability are due to effects upon the salt linkage. At *p*H values less than 3·0 and greater than 11·0 and to some limited extent between 3·0 and 11·0, there are changes taking place relative to the short linkages. The effect upon the hydrogen bond type of linkage is more in evidence in the zones *p*H 1·0-3·0 and *p*H 12·0-14·0 as can be seen from the sharp decrease in shrinkage temperature. This sharp change in structural stability is indicative of a more drastic change thah the mere back titration of charged groups since a goodly number of these groups have already been weakened or broken in the less acid or alkaline zones. The writer believes that acid in the very low *p*H range or alkali in the very high *p*H range has a real and pronounced effect upon the short link or hydrogen bond structure.

Further in the *p*H range 2·0-3·0, a large amount of swelling (curve A) obtains, which undoubtedly affects the ability of the structure CO—NH to maintain itself and thereby structural weakness occurs. At *p*H 1·0 swelling again reaches a rather low value and though osmotic pressure has decreased, the reduction in cohesion is still evident. Braybrooks *et al.* pointed this point out in an earlier study.

A careful study of curves A and A′ of Fig. 1 indicate the direct relation of the swelling and hydrothermal stability of collagen. In the *p*H zone 5·0-10·0 swelling is at a minimum while the shrinkage temperature is maintained at maximum values. As swelling increases (*p*H values less than 3·0 and greater than 10·0) shrinkage temperature decreases.

Curve B′ of Fig. 1 shows the shrinkage temperature of collagen treated with acid-salt and alkali-salt solution over the *p*H range 1·0 to 13·0. All solution were 2·0 N with respect to NaCl. These data show that the salt has a drastic effect upon both the swelling and the shrinkage temperature of the collagen. It is to be noted that the salt plays a role only in the *p*H range less than 9·0. In this zone the shrinkage temperature is much greater than when no salt is present and the swelling is drastically decreased. A careful study of the four curves indicates the following trends: (*a*) at *p*H values greater than 10·0 salt has little or no effect upon swelling and thus decrease in hydrothermal stability results; and (*b*) at *p*H values less than 10·0 reduction in swelling makes for a closer fibril structure and thus causes a strengthening of the cohesive forces even in the very acid zone.

It is apparent that the reduction in swelling and thus dehydration of the collagen in the *p*H range 0·5 to 4·0 causes a great increase of hydrothermal stability which must be largely attributed to the strengthening of the hydrogen bonds in this *p*H range.

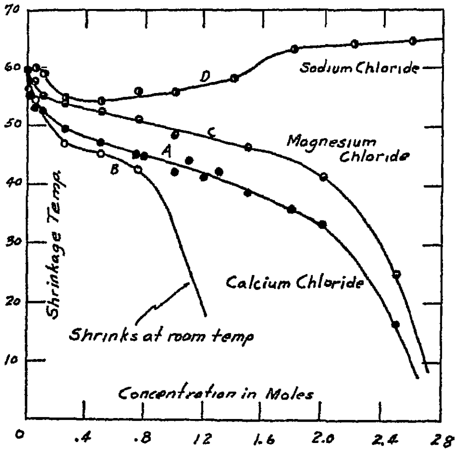


FIG. 2.

Such facts as given bear out the statement that dehydration within reason causes greater structural stability, whereas swelling has the reverse effect. The data given in Fig. 1 substantiates the contention that it is the effect upon the short linkages in the main rather than upon salt linkages that brings about structural weakness. This conclusion is substantiated by the studies of Braybrooks *et al.* and by those of Jordan Lloyd.

Figures 2 and 3 show the hydrothermal stability and swelling of proteins in various salt solutions, such solutions ranging in molal concentrations from 0 to 2·8. These data show that dilute solutions of NaCl at *p*H 5·5 cause a reduction in structural cohesive forces of collagen—an action undoubtedly upon the zwitter ions or upon salt bridges. Strong solutions of NaCl, at this same *p*H value due to dehydration forces, cause decreased swelling with subsequent increased hydrothermal stability.

Many investigators have found that collagen and even tanned leather shrinks in concentrated solutions of $CaCl_2$ at normal temperatures. Curves A and B of Fig. 2 show that even dilute solutions of $CaCl_2$ cause reduced structural cohesion and that more concentrated solutions so weaken the intermolecular structural network that collagen collapses upon itself. It therefore is evident that $CaCl_2$ affects both the hydrogen bond and the salt-type linkage. In very dilute solution, this effect is upon the salt linkage and apparently is much more effective than the monovalent salt. It must be remembered that $CaCl_2$ actually causes swelling in the isoelectric zone and has a positive peptising action upon collagen. Thomas [15] showed that the degree of hydrolysis of collagen by $CaCl_2$ was three times greater than with NaCl. In addition $CaCl_2$ does not have a dehydration effect but has an actual swelling one and thus must exert a positive strain upon both the short and salt linkages. At concentrations greater than 0·8 mole, this salt must exert its main activity upon the hydrogen bonds since almost complete structural breakdown obtains. Magnesium chloride exerts the same general trend as the calcium salt as can be seen from curve C of Fig. 2. A close study of the curves of Fig. 2 and 3 indicate in a rather positive manner the relation of swelling to hydrothermal stability of collagen.

The investigations so far discussed have been those for acid, base or salt treated collagen. The work next to be described will be that of formaldehyde-treated collagen. Fig. 4 shows swelling and shrinkage temperature data for native and formaldehyde-treated collagen. These data show rather conclusively that for formaldehyde-treated collagen, the shrinkage temperature is in line with the swelling data. In the acid range *p*H 1·0-3·0 little formaldehyde is found, rather high swelling obtains and as a consequence low shrinkage temperature results. At *p*H values greater than 3·5 swelling is greatly reduced by the action of the formaldehyde upon the collagen and increased shrinkage temperature results both because of the reduced swelling and aldehyde fixation. In the alkaline zone (*p*H 8·0-12·0) swelling is little greater than that obtaining in the acid region and thus the increased hydrothermal stability remains essentially constant as can be seen from Fig. 4.

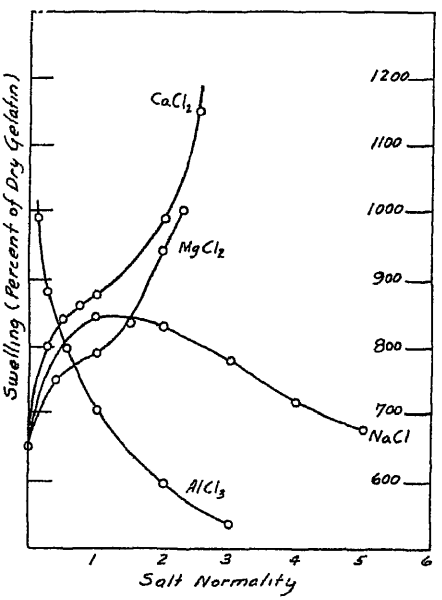


FIG. 3.

De-aminisation of collagen through the action of nitrous acid must of necessity destroy a number of the electrovalent linkages existing between adjacent polypeptide chains. Such internal changes would then evidence themselves in the breakdown of internal cohesional forces. This particular type of structural change should be especially marked in the so-called *p*H stability zone. Data taken by Theis and Lams [16] relative to hydrothermal stability and swelling of de-aminised collagen over the *p*H range 1·0-12·0 show definitely that in the region of breakdown of such salt linkages, because of the nitrous acid action, namely at *p*H values greater than 6·0, there is a sharp increase in swelling and a sharp breakdown in internal cohesional forces of the de-aminised collagen. A comparison of data for normal collagen shows that in the acid zone, *p*H 3·0-5·0, the shrinkage temperature of the de-aminised collagen is just a little greater than that for the normal collagen and can undoubtedly be accounted for by the decreased swelling obtaining for the de-aminised protein. In the alkaline zone, pH 6·0 -12·0, the shrinkage temperature shows a sharp decrease in the specific regions in which histidine and lysine would normally have played an important role. That degree of swelling and hydrothermal stability are interrelated is indicated by the great swelling of de-aminised collagen taking place in the alkaline zone which undoubtedly is responsible for the weakening of the internal cohesive forces of the nitrous acid treated collagen.

---

[15] Thomas and Foster, *Ind. Eng. Chem.*, 1925, 17, 1162.

[16] Theis and Lams, *Chemistry of Leather Manufacture* (Reinhold, New York, 1945).

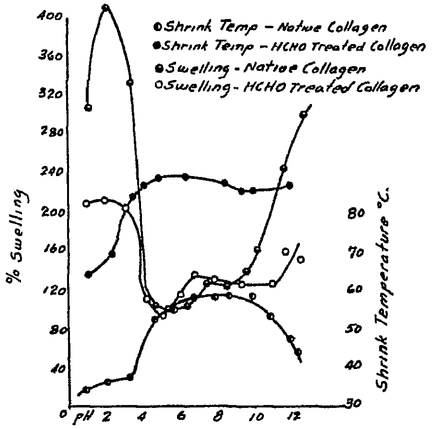


FIG. 4.

Summary.

Theis *et al.* have repeatedly stated that the swelling and the hydrothermal stability of collagen must of necessity be interrelated. This statement is very well evidenced by a careful study of their investigations relative to the effect of acids, bases, salts and fixation agents upon collagen. In these studies, it is shown that in the isoelectric or buffer zone there are two reactions : (*a*) the main one an action upon the salt-like linkages, and (*b*) an action upon the short links or hydrogen bonds thus changing the internal structural cohesive forces of the collagen. These investigations have indicated that at pH values less than 3·0 or greater than 10·0, the action of the acid or base is upon the hydrogen bond linkages since there is a sharp break in the hydrothermal stability of the collagen in both these zones.

The data relative to the action of mono- and divalent salts has been discussed and it has been shown that such salts as calcium and magnesium chlorides have an entirely different effect upon the structural cohesional forces of the treated normal collagen.

Biochemistry Laboratories,
Lehigh University, Bethlehem, Pa.

*GENERAL DISCUSSION**

Dr. J. L. Stoves (*London*) said : A study of the relations between cross-linkage reactivity and the supercontraction of keratin fibres has given results of interest in connection with the swelling of protein fibres. If fibres of human hair and horse hair in equilibrium with air of 65 % R.H. at 22·2° C., are heated at 100° in dry *m*-cresol, then on drying such fibres, they contract in length to the extent shown below :

Time (hr.) . . .	1·0	2·5	4·5	6·0	7·5	10·0
% Contraction—						
(*a*) Human hair .	2·8	3·7	4·6	4·6	4·6	4·6
(*b*) Horse hair . .	1·2	7·2	12·1	13·3	13·3	13·3

---

* On two preceding papers.

The outstanding feature of this experiment, as well as that in which *m*-cresol plus 10 % water was used, is the very large difference in behaviour of the two types of fibre, yet both are classed as keratin fibres. Many other examples of wide differences in behaviour of fibres in the keratin group have been encountered, e.g., human hair treated for 15 min. at room temperature with Cl_2 water, followed by 48 hr. washing, contracts in length by 33 % on drying. Under similar conditions horse hair diminishes in length by only 2 %. These effects do not seem to be due to gross differences in histological structure, e.g. thick cuticle or excessive medullation, characteristics known to account for differences in behaviour of primary hairs from certain mammals.[1] Variation in the nature and number of cross-linkages in the keratin macro-molecules of the two types of hair is a possible explanation of the result. A further possibility which should be considered, however, is variation in the region intermediate between the microscopic and molecular, i.e. dimensions explored by electron microscopy. The following observation shows the desirability of further work along these lines. After 4 hr. in 98 % formic acid at 100° horse hair contracted in length by 12 %. The exceptional result, however, was that these fibres began to split longitudinally into a number of discrete fibrils. Mechanical treatment ultimately caused the fibrils to break up into spindles. Human hair did not behave in this way. Degradation of the fibre is readily understandable. The normal, non-medullated keratin fibre, however, is considered to consist of a mass of spindle-shaped cortical cells cemented together by some intercellular protein, the whole being surrounded by cuticle scales. Such a structure would be expected to break up into an irregular mass of spindles, rather than into a number of fibrils running the whole length of the original fibre. The results suggest that in horse hair there may be a marked columnar organisation of biological cells in the cortex.

Dr. G. Salomon (*Delft*) (*communicated*) : The experiments of Miss Lloyd illustrated the close relationship between fibres and rubber. The rubber-like state is reached by weakening the intermolecular attraction through the action of solvents. Very similar phenomena can be observed by introducing either the strongly polar nitrile group in certain co-polymers of irregular structure or the moderately polar chlorine atom in the regular pattern of the natural rubber chain. The following examples will illustrate this line of thought.

The co-polymers from butadiene and acrylic nitrile are rubbers, but the corresponding co-polymers from dimethylbutadiene and acrylic nitrile become rubber-like only above 100°.[2] They can be extended in boiling water and such a stretched strip will contract at room temperature only under the action of swelling agents. The process of contraction precedes swelling and the phenomenon resembles very much Fig. 2 of Miss Lloyd's paper.

The well-known hydrochloride from natural rubber[3] can be conveniently stretched in hot water (using slightly vulcanised strips) and here also swelling at room temperature in organic solvents is preceded by an axial contraction. Our conclusion that the fibre form is in this case due to the regularity of the chain is based on the observation (unpublished) that the hydrochloride from polyisoprene, which lacks the regular chain configuration, is still elastic at room temperature.

Mr. L. Peters (*Leeds*) said : In this paper the swelling and shrinkage temperature of collagen at various *p*H values are interpreted by Prof. Theis as indicating a breakdown of hydrogen bonds between peptide links of adjacent main chains. Presumably, the effect of acid in breaking the hydrogen bonds is due to combination of protons with the peptide group according to the equation

---

[1] Stoves, *Nature*, 1943, **151**, 304 ; 1946, **157**, 230 ; *Proc. Roy. Soc. Edin. B*, 1945, **62**, 132.

[2] Salomon and Koningsberger, *J. Polymer Sci.* (in print).

[3] Gehman, Field and Dinsmore, *Proc. Rubber Techn. Conf.*, 1938.

$$—\mathrm{CO}\,.\,\mathrm{NH}— + \mathrm{H}^+ = —\mathrm{CO}\,.\,\mathrm{NH}_2{}^+—,$$

or by some such similar mechanism. The ionisation of the resulting acid can be characterised by a dissociation constant whose magnitude will be of the same order as that of similar compounds such as acetamide or the polyamide Nylon. The pK_a of acetamide is approximately 0·5, and the data for the combination of Nylon with sulphuric acid[4] enable a calculation of the strength of this " peptide acid " to be made giving $pK_a = -1{\cdot}1$.

There is some evidence that proteins combine with more acid at *p*H values below 0·5 and if this is not due to hydrolysis of the peptide link it could easily be explained as combination according to the above scheme. But with the large value which the dissociation constant must have, such combination must be extremely insignificant in the range of *p*H = 1·0 to 3·0.

It is well-known that the swelling on the acid side can be attributed to a Donnan membrane effect caused by the combination of protons with the carboxylic ions on the protein and this was the basis of the successful Procter - Wilson theory. Large salt concentrations, by " swamping " the membrane effect, will cause the " internal " *p*H to decrease to a value approaching that of the external solution. This should lead to greater combination of acid with the peptide bonds at any given *p*H and, if the theory put forward in this paper were correct, to a greater decrease in the shrinkage temperature. Fig. 1 shows, on the contrary, that in the presence of salt there is a greater hydrothermal stability than in the absence of salt. For these reasons it appears that the phenomena cannot be explained in terms of breakdown of hydrogen bonds.

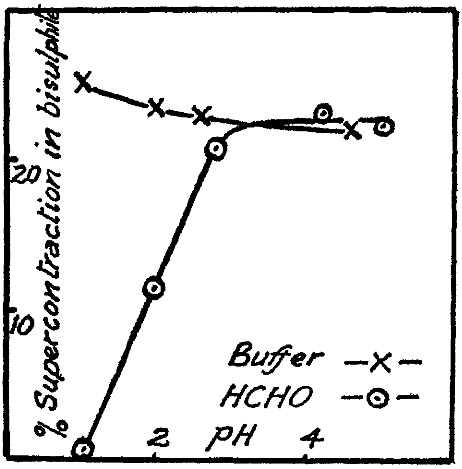


FIG. 1.—% Supercontraction in bisulphite.

Even if all the salt links are back-titrated at the *p*H values concerned, the Donnan effect due to these groups would still exist and be " swamped " by the increase in concentration of diffusible ions. If " dehydration can cause a great increase in hydrothermal stability " as stated by Prof. Theis then the suppression of swelling by salt can explain the above anomaly without the need for involking hydrogen bonds.

Finally, it must be emphasised that whilst the COO^- groups are *almost* completely in the —COOH form at *p*H 2·0, the breaking of the *last* links in a structure may be the decisive factor.

Dr. J. L. Stoves (*London*) said : Allied to the hydrothermal shrinkage of normal and HCHO-treated collagen is the contraction of normal and HCHO-treated keratin fibres in boiling solutions of bisulphites. Human hair fibres boiled for 1 hr. in 2 % HCHO solutions of *p*H 1-5, supercontract in boiling 5 % sodium bisulphite as shown below.

Now the amount of HCHO fixed by keratin at *p*H 1 is relatively small and the stress-strain curves of treated fibres show little difference from those of fibres boiled in buffer alone. Yet, within the experimental error, supercontraction is completely inhibited by HCHO treatment. Furthermore, if amino-, acid amide and guanidyl-groups are removed from fibres, by treatment with van Slyke reagent in the presence of 2 N. HCl, prior to HCHO treatment, the following results are obtained :

---

[4] Boulton, *J. Soc. Dyers Col.*, 1946, **62**, 65.

Time in HNO_2 (hr.)	0	5	16	24	48	96
% Contraction after $NaHSO_3$—						
(*a*) HCHO *p*H 1	2	9	17	35	37	38
(*b*) Buffer *p*H 1	25	33	—	48	55	68

Increasing modification of side chains leads to increasing supercontraction of the buffer-treated fibres, yet throughout the series the HCHO-treated fibres have a supercontraction some 20 % less than that of the buffer fibres. Stabilising linkages can be formed, therefore, in keratin fibres from which amino-, amide and guanidyl-groups have been removed. If these linkages are formed by reaction with imino-groups of the main polypeptide chains it is difficult to see why much more HCHO is not fixed by keratin in the acid range. Furthermore, treatment of hair with HCHO in the isoelectric region, where fibre swelling is a minimum, would be expected to give rise to a maximum number of stable linkages. In actual fact few if any such linkages are formed. More recently, the contraction of HCHO treated fibres in *m*-cresol plus 10 % water at 100° C. has been examined. The results show that here again, HCHO treatment increases the stability of the fibres.

It has been suggested that small amounts of thiol, formed by reduction or by hydrolysis of the cystine linkage (the latter was shown to occur by Shinohara), may be involved in the formation of stabilising acid linkages (—S—CH_2—S—).[5] The appearance and contraction in length after bisulphite treatment of fibres treated with HCHO of *p*H 1 and buffer of *p*H 1 respectively, followed by boiling in dilute acid, leave no doubt that HCHO-treated fibres contain new linkages which resist acid hydrolysis.[6]

In acid solutions only a small number of stable linkages would be expected to be formed by reaction of HCHO with —SH groups, which agrees with the low fixation of HCHO. Yet small numbers of stable linkages formed throughout the polymeric system may exercise a considerable effect on supercontraction. This view is not unreasonable in the light of Staudinger's work on the solubility and swelling of polystyrene cross-linked with *p*-divinyl benzene.

Prof. J. D. Bernal (*London*) said : I would like to draw the attention of the Conference to the remarkable work carried out particularly by Schmitt, Hall and Jakus[7] on collagen and other natural fibres. The banded structure of collagen, with a periodicity of 70 A. had been known for some time, but in this work the higher resolving power showed the existence of five sub-bands of markedly different intensity. These can be inferred from the X-ray photographs through the relative intensities of the different orders of reflection of the meridianal spacing, but were directly shown in the electron microscope. Two features of particular interest were apparent. The arrangement of sub-bands was unsymmetrical; i.e. the fibres were unlike viewed from one end or the other, indicating that the collagen fibre was produced continuously from one end by a rhythmical process with a complex rhythm. Further light on the nature of this process was shown by experiments on the staining of the different bands by various reagents, which indicated that each band consisted for the most part of a different amino-acid. This would lead us to a somewhat different conception of fibre protein molecules than has hitherto prevailed, namely that the amino-acid residues of like character may succeed each other in groups of ten or twenty, rather than the intimate interspersal originally put forward in the hypothesis of Bergmann and Niemann. Collagen itself does not seem at present to have shown any natural unit at right angles to the bands, so that these observations have not as yet thrown light on molecular swelling and shrinkage in this direction. Some photographs, however, show individual fibres elongated by accidental tensions in the specimen by as much as five times, and the band structure is also stretched out, though there would appear to be somewhat more stretching between the bands than in them. All these observations are somewhat difficult to reconcile with any of the chain hypothesis of protein structure that have been put forward, because if a molecule can be stretched five times it must have been far more curled in the first place than has been accepted.

---

[5] Stoves, *Trans. Faraday Soc.*, 1943, **39**, 294.

[6] Stoves (in press).

[7] Hall, Jakus and Schmitt, *J. Appl. Physics*, 1945, **16**, 459.

SWELLING AND CYTOPLASMIC VACUOLATION.

By E. Fauré-Fremiet

Received 15th July, 1946.

Osmotic swelling of numerous animal and plant cells in hypotonic medium may lead to vacuolar separation of protoplasmic fluids.[1]

Between certain limits of hydration and of temperature, this osmotic swelling is reversible;[2] in some cases, however (Leicht[3] with *Echinometra* egg) we observe, when the cellular volume has exceeded a certain value, a slight contraction shortly followed by a further increase coinciding with the formation of vacuoles; this new state is irreversible and of pathological order.

Some chemical actions (treatment with acid or alkali) affect the imbibition of water by the cytoplasm and may lead to the formation of vacuoles. Giersberg[4] and Rumjantzeff[5] suppose that the amount of water entering the cell under the effect of different ions may exceed the binding power of cytoplasmic colloids and bring about the separation of an aqueous vacuolar phase. However, Speck[6] as well as Heilbrunn[7] point out that this apparent separation represents the increase of an aqueous phase pre-existing in very finely dispersed state.

It is likely that one can find in the case of living cells the two mechanisms which Jordan Lloyd recognises, in the case of proteins in general and especially of leather, under the respective terms "plumping" (or the Hofmeister effect), and swelling, or the Donnan effect.[8] The intake of water brought about by the Donnan effect may exceed the binding power of non-diffusible colloids for water, and produce separation of a protein coacervate and of its equilibrium fluid.

It is known that ammonia and its salts penetrate swiftly into living cells, as also do a certain number of amines. Garrault subjects *Paramecia,* previously treated with congo red and neutral red, to dilute solutions of amines; the acidity of digestive vacuoles is soon neutralised and the colorisation due to neutral red turns to yellow, an effect which may be obtained *in vitro* with the same compounds in the neighbourhood of *p*H 7·8.

---

[1] Anitschow, *Verhandl. deutsch. Pathol. Ges.*, 1914, **17**, 100. Lepeschkine *Biologia generalis*, 1925, **1**, 368.

[2] Lucke and McCutcheon, *Arch. Pathol. Lab. Med.*, 1926, **2**, 846.

[3] Leicht, *Proc. Soc. Exper. Biol. Med.*, 1935, **33**, 207.

[4] Giersberg, *Arch. Ernwicklungsmech.*, 1922, **51**, 150.

[5] Rumjantzeff, *Arch. exper. Zellforsch.*, 1926, **3**, 115.

[6] Spek, *Kolloidchem. Beihefte*, 1920, **9**, 259; *Zeitschr. Zellf. Gewebelehre*, 1925, **1**, 278.

[7] Heilbrunn, *The Colloid Chemistry of Protoplasm*, Berlin (1928).

[8] Jordan-Lloyd, *The Swelling of Proteins*, London, 1933, 74, 111.

Anna Litvac * in the early phases of a study on the effects of different amines upon fibroblasts grown *in vitro* has shown a very interesting fact: methylamine, ethylamine, phenylethylamine and tyramine, added to the cultures at concentration of about 0·006 %, bring about a strong cytoplasmic vacuolation, readily visible in all the cells of the area of growth; the diamines such as histamine and putrescine, on the other hand, fail to produce this phenomenon.

With H. Garrault[9] we have resumed these experiments on different ciliates and have shown that *only* the alkylamines (methyl-, trimethyl- ethyl-, propyl-, *iso*-amylamine) produce vacuolation at concentrations ranging from 0·01 to 0·1 %, while diamines, arylamines and so on give negative results, or produce cytoplasmic change At first sight, the length of the carbon chain seems to play a small part, and the necessary structural characteristic is the presence of a terminal polar group (lyophilic) and of another non-polar group (hydrophobic).

New research, of a volumetric order, done with *Paramecium* has shown that protoplasmic swelling produced by penetration of alkylamines is very slight and cannot be accurately measured; the very numerous newly formed vacuoles, represent a volume clearly greater than that of the water taken up by the cell; one must conclude, therefore, that the penetration of a small amount of alkylamine provokes separation not only of the water which brings about swelling but also of a part of the normal water in the cell.

The interpretation formerly put forward remains.[8] We assume that the alkylamines become fixed on the cytoplasmic constituents by their polar groups and because of their non-polar free groups decrease the affinity of the whole for water; separation leading to vacuolation thus appears like a case of complex coacervation (Bungenberg de Jong). This interpretation is not necessarily a general one. P. Dustin, Jr.,[10] obtained as a matter of fact, vacuole formation in many sorts of red blood cells with all the amines, and not only with those possessing a terminal hydrophobic group. Recalling Chlopin's and Kedrowski's conceptions, P. Dustin assumes that amines combine with cytoplasmic basophilic substances, such as fine nucleoproteinic granules, and increase their affinity for water to such an extent that they dissolve. With H. Garrault we have confirmed Dustin's observation on red blood cells and on larger monocytes of frog blood.

In so far as the ciliates are concerned, it should be noted that all do not react in the same way to the alkylamines and that *Dileptrus,* for instance, never show vacuolar separation; in that case, however, it was impossible to follow penetration of the amines and, if we increase its concentration in the medium, we bring about an alteration of the superficial film followed by cytoplasmic change.

It thus appears that swelling phenomena and separation brought about by amines depend on the one hand on the molecular structure of these substances, and upon the other on the physicochemical properties of the cells on which they are caused to act.

College de France,
Laboratoire d'Embryogenie Comparee,
Paris, France.

---

* Shot by the Nazis in the Auschwitz Camp, August, 1942.

[9] Fauré-Fremiet, *Bull. Soc. Chim. Biol.*, 1944, **27**, 96.

[10] Dustin, P. (Jn.), *Arch. Biol.*, 1944, **55**, 285.

ON THE SWELLING OF STARCH GRAINS.

By N. P. Badenhuizen.

Received 10*th July*, 1946.

A few years ago I concluded that the argument that two chemically defined substances exist in the starch grain was untenable. Since then the work of K. H. Meyer,[1] has shown that most of the starch molecules are ramified (the amylopectine) and are mixed up by the single chains of amylose. Both are built up from the same chemical units. Nevertheless, amylopectine molecules have varying degrees of ramification. Likewise, the amylose consists of longer and shorter chains. During the swelling of the starch grain the amylose molecules go into solution first, some of them remaining in the meshes of the amylopectine. On the other hand, a few ramified molecules may already appear in the solution. Therefore, the names " amylose " and " amylopectine " only serve the practical end of distinguishing two groups of molecules of different form. We can say with Bourquelot,[2] that the starch grain contains not two, but a great many, kinds of carbohydrates, which have the same origin, and develop in course of time into a series with different resistance. Together they form a reticular disperse system, which crystallises after the intake of water.[3]

This conception solves all the controversies as to the isolation of the two alleged constituents. Moreover, it is consistent with the chemical and microscopical observations.[4]

We always see the starch grain in optical cross-section consisting of a system of layers, in each of which we can distinguish lamellæ. The birefringence of one layer decreases from within outward. The periodic supply of carbohydrates is responsible for this structure. The amyloplast builds large molecules out of sugars, with the aid of enzymes like the P- and Q-enzymes.[5] At a certain concentration the starch molecules separate from the stroma. They are deposited in a more or less radial direction, corresponding to the specific structure of the amyloplast, and because water molecules are partly withdrawn. After new molecules have taken the place of the old ones by diffusion, the process is repeated. In this manner the lamellæ are formed and these must be very closely connected together.

However, the supply of carbohydrates to the amyloplast does not keep step with the rate of precipitation, and so the separation becomes more and more incomplete, the molecules being smaller and ramified to a less extent, the chains shorter and more hydrated, their connection looser. When this process comes to a stop below a certain concentration of carbohydrates, the formation of a layer has been completed.[6] One kind of amyloplast (e.g., potato starch) allows this critical concentration to be lower than another kind (e.g., tapioca starch). In the first case the connection between the layers is relatively loose, in the second case, the connection is more intimate, and the starch grain becomes more homogeneously built in the radial direction.

---

[1] Meyer, *Adv. Coll. Sci.*, 1942, **1**, 143.

[2] Bourquelot, *Compt. Rend.*, 1887, **104**, 177.

[3] Speich, *Ber. Schweiz. Bot. Ges.*, 1942, **52**, 175.

[4] Badenhuizen, *Rec. Trav. bot. néderl.*, 1938, **35**, 559.

[5] Bourne and Peat, *J. Chem. Soc.*, 1945, **11**, 877, 882.

[6] Badenhuizen, *Prot.*, 1939, **33**, 440 ; *Chron. Bot.*, 1939, **5**, 199.

The above conceptions accord fully with microscopical observation. Combined with the present conception of the chemical structure they furnish the clue to the understanding of all phenomena the starch grain may display while swelling. At the same time they put the starch grain before us as a biological object, and explain how it is that the plant succeeds in storing material, of which it can dispose at any moment.

The above depicted mode of growth is called apposition, and is based on the observation, that all the layers in the starch grain are similar, no special membrane surrounding the grain, as I have shown.[1] Some people still believe in the existence of an outer membrane which is supposed to have all sorts of properties to explain different phenomena, but these introduce many discrepancies and unfounded assumptions. One author suggests that the alleged membrane is thinnest in small starch grains,[7] another accepts the reverse,[8] and a third goes as far as to assume a different membrane for each kind of starch grain.[9]

When we remove one or more of the outer layers, the remaining part develops a vesicle in the usual way after gelatinisation. Canna starch grains, brought into 20 % chromic acid, show swelling of the outmost layer, which separates and dissolves into the liquid ; the next layer then undergoes the same fate, and so on. After heating in dry condition (formation of pyrodextrins) potato starch grains develop a system of separated layers, or better, vesicles. All the layers react in exactly the same way. We can open a grain with a micro-knife, the inner mass does not dissolve. There is no protecting membrane. In fact, the wall of the vesicle, produced by the swelling of the starch grain, is built by the confluence of all layers, which we can distinguish clearly after limited gelatinisation. In certain circumstances we can separate them again.

The lamellæ of one layer always swell together, even if we separate the layers, e.g., after bringing pyrodextrin from potato starch into cold water. In this case it is clear that the lamellæ of low refractive index are less resistant than those with high refractive index. Their molecules are the first to dissolve. All chains have been broken down by the action of heat, but the network in the more resistant layers is still strong enough to retain its swelling power. The molecular structure is loosened by shortening of the chains and lateral dispersion, so that the substance now swells in cold water. The more the starch grains have been heated, the more readily does this process go, till the molecules are so far broken down that they have all become soluble. At the same time the iodine colour changes from violet to red.[10] Tapioca starch grains do not separate the layers after treatment in the above mentioned way. Here the layers remain firmly connected and each develops in cold water one single vesicle in the usual way.

When we add a little iodine to the preparation of roasted potato starch in water, we see a blue-coloured substance streaming forth while the layers swell and separate. This cannot be due to an internal osmotic pressure. The layers separate by tangential swelling only. In a spherical starch grain the surface of the various layers increases proportionally to the square of the radius. Every surrounding layer contains more molecules than the underlying one, assuming that the layers all show the same morphological structure, as is very probable from microscopical observations. Moreover, by comparison of small and large starch grains of one kind, we find no difference in chemical composition,[11] while both show the same X-ray diagrams.[4] Here we have an explanation for the increase in swelling capacity of the layers from within outward, which makes them separate.

---

[7] Höppler, *Kolloid Z.*, 1942, **101**, 305.

[8] Wieler, *Prot.*, 1943, **37**, 171.

[9] Knjaginichev, *Compt. Rend.* (*Doklady*) *Ac. Sci. U.S.S.R.*, 1939, **23**, 92.

[10] Badenhuizen and Katz, *Z. physik. Chem.*, *A.*, 1938, **182**, 73.

[11] Stamberg and Bailey, *Cereal Chem.*, 1940, **17**, 309, 319.

Furthermore, the difference in surface will be greater between two outer layers than between two layers near the centre. Therefore, we often see a separation of the peripheral layers, leaving an intact nucleus behind.

As a result immature small grains give the impression that they are more homogeneously built than the larger ones. The same holds true for the inner mass of a large grain, as compared with the periphery. Likewise large grains are the first to be attacked by a solution causing swelling, because the swelling pressure of their outer layers is greater than that of the outer layers of a small grain. This explains the fact that starch which has been slowly heated in water has no sharp point of gelatinisation, but swells within a certain range of temperature. When we count the number of grains swelling within short ranges of temperature (e.g., 2° c.), we can plot a curve of probabilities, the arithmetical mean indicating the most exact measure for the gelatinisation.

The following experiments prove that the swelling of the starch grain is only due to tangential swelling of the layers due to their peculiar structure responding to the altered circumstances in the surrounding liquid, the osmotic pressure of the freed molecules being of no importance.

We cut off the posterior part of an eccentric starch grain (e.g., potato) and bring it with the aid of an eye-shaped micro-needle into a droplet of 1 % NaOH. The pieces swell and form a cup-shaped vesicle with an entrance, through which we can insert a needle. The contents, which consist of the solvated molecules, come out by slow diffusion. In an earlier paper [4] I criticised the experiments of Hess and Rabinowitsch,[12] which seem to show the contrary.

We imbed starch grains in gum-arabic. From the hard mass thin scales are cut with a razor. After solution of the gum in water the preparation contains many starch grains, surrounded by beautiful rings made up of semilunar elements.[4, 13] A study of those rings shows that they originate from the outer parts of the grain, which have been damaged by the knife and now swell in cold water (see below). Because they remain attached to the inner intact mass of the grain, they are laid into (often very regular) folds, and so give a good demonstration for the tangential swelling.

We can now fully understand the swelling of the starch grain. All layers swell at the same time. But as the outer layers show a stronger swelling power in the tangential direction than the inmost layers, and because they all stick together, the inner part is stretched out by the outer part. In consequence a small cavity can be detected in the centre of the grain. This cavity enlarges as swelling proceeds, the volume of the grain increases, the thickness of the wall of the growing vesicle decreases. The constituents of the inner layers being most subject to dispersion must send out their molecules into the interior space of the vesicle. So long as the wall of the vesicle is smooth, it is very elastic, which is in accordance with the conception of 3-dimensional giant molecules, held together by network cross linkages.[1]

A certain force is necessary to break the lateral cohesive forces. This causes the swelling of the starch grain to start suddenly. We can speak with Katz of a " grown structure ", which has to be destroyed first.[13] This " grown structure ", however, is not a complicated system of small cubic units.[14] It often appears at the beginning of the swelling as a result of fine radial fissures, cutting off small pieces from the resistant layers. Their origin is secondary and artificial.

The starch grain may be plastically deformed by pressure without disturbing its coherence. It then swells unsymmetrically after heating in water, according to the local deformation, and dependent only on the way the molecules are arranged.

---

[12] Hess and Rabinowitsch, *Kolloid Z.*, 1933, 64, 257.

[13] Badenhuizen, *Prot.*, 1937, 28, 293.

[14] Badenhuizen, *ibid.*, 1937, 29, 246.

As soon as molecule-ends are set free after damaging the grain, they attract water molecules and the damaged part swells in cold water. It is very instructive to cut a starch grain slowly with a micro-knife in a droplet of water containing congo red. Where the knife went through the grain, the substance colours red, an indication that it is swollen. Lepeschkin's [15] assumption of a heat-effect caused by friction, and his theory as to that the amylose and amylopectine result from chemical reaction between water and starch substance is not satisfactory.

The amylose shows the phenomenon of ageing after extraction from the grain and subsequent purification.[16] By gradual association of the particles the substance becomes insoluble. We can also observe similar changes in the starch under the microscope. After slow and partial gelatinisation of starch grains (e.g., Canna) we observe large wedge-shaped fissures in the grain (point outward). The substance between those fissures is likewise wedge-shaped (point inward), and shows a remarkably increased resistance during the further stages of swelling. In other cases, too, the internal mass, which should disappear first during swelling, remains isolated and possesses an abnormal resistance.

From micro-dissections [4, 17] it appears that the dry grain (e.g., of potato) is very hard. As soon as it is surrounded by water the substance may be cut easily, the elasticity increasing gradually during the process of gelatinisation, reaching a maximum and decreasing again after the vesicle folds up and is finally dispersed into the liquid. Wheat or rye starch grains are not so elastic, as is the case also with starch grains of potato when swollen in pyridine or phosphoric acid. It is obvious that the degree of elasticity is determined by the intake of water into the spaces of the molecular network, but also by the size of the chains and the weakening of the lateral cohesive forces. The vesicles of pyrodextrins are much less elastic than those of gelatinised natural starch, while vesicles which have been exposed to the influence of bacteria, or which have been impregnated with iodine, show no elasticity at all.

The following observation is very striking; it demonstrates the strong attractive power starch substance exerts on water molecules by gradual evaporation of the surrounding water. As soon as a certain limited quantity of water has been left in the micro-droplet, the grain absorbs the remaining water energetically, with considerable softening of the starch substance. Lintnerised grains, which before are copper-coloured in iodine, and which are very brittle, now stain a deep blue and swell, while its substance is softened. On addition of fresh water they may dissolve into the liquid. Stains, which do not penetrate the lintnerised grain when a sufficient quantity of water is present, are sucked in during the last moments of drying. Vesicles of pyrodextrins lose the elasticity completely and even disperse.

This change in resistance is reversible only in the natural starch grain. The degraded products are irreversibly broken down. Fischer [18] observed similar phenomena for sphærites of $CaCO_3$ and inulin. Probably the surface tension of the water plays an important part, forcing the water molecules into the starch substance, widening the meshes of the network, and eventually rupturing linkages. The last-mentioned process is irreversible.

We have pictured here the structure and swelling of the starch grain, so far as present knowledge permits. Several problems remain to be solved, but a general insight has undoubtedly been obtained. We have good hope that future experiments will lead to a final conclusion and a deeper understanding of the details. An exact knowledge of the starch grain is of fundamental importance for the solution of technical problems. With the amyloplast it takes us straight into the problem of life.

---

[15] Lepeschkin, *Kolloid Z.*, 1942, **101**, 196.
[16] Meyer, Bernfeld and Wolff, *Helv. Chim. Acta*, 1940, **23**, 854.
[17] Badenhuizen, *Chron. Bot.*, 1938, **4**, 123.
[18] Fischer, *Cohn's Beitr. Biol. Pfl.*, 1902, **8**, 53.

Botanical Institute,
Amsterdam.

THE SWELLING AND SHRINKING OF PROTOPLASM.

By W. Seifriz.

Received 10*th July*, 1946.

The control of water content is one of the most fundamental properties of living matter and one of the least well understood. As osmosis, the phenomenon is quite clear, but there is much evidence to indicate that water control in protoplasm is not just a simple case of osmosis. For example, if the leaves of the sensitive plant, *Mimosa*, are irritated, they close immediately through collapse of the cells of the pulvinus due to exosmosis of water; but there has been no rapid change in the relative concentration of osmotically active substances in and out of the cell; a sudden change in relative temperatures which would permit thermal osmosis, likewise could not have taken place. One possibility remains, a change in relative potentials; this would account for electro-exosmosis, but there is no evidence for it. The one-way flow of water through the intestine wall is another example of water control which is not readily interpreted. Here, again, electro-endosmosis is the most plausible explanation if we view the phenomenon as osmotic.

The mechanism by means of which gelatin soaks up water is that of imbibition and always involves pressure. Formerly, osmosis and imbibition were regarded as wholly distinct phenomena, but the introduction of the Donnan [1] equilibrium into our ways of thinking about the uptake of water by colloidal systems, caused certain investigators [2] to regard osmosis and imbibition as identical phenomena, certainly in so far as the forces involved are concerned. J. Loeb [3] said, " the reason that osmotic pressure, the viscosity of protein solutions, and the swelling of protein gels are all influenced in a similar way by electrolytes is that all three properties are, in the last analysis, functions of one and the same property, namely, osmotic pressure." This may be true, but an attempt will be made to show why some think it not true, why some regard imbibition as a process quite distinct from osmosis. The latter term has heretofore always been applied to the force involved in a system where a membrane encloses a solution. Freundlich [4] states emphatically that it is fundamentally false to regard osmotic pressure as identical with imbibition pressure, for there is heat of imbibition, and not heat of osmosis. This may, however, be merely a matter of degree. More significant is the fact that osmotic pressure is proportional to concentration, whereas imbibition pressure is proportional to a high power of the concentration.

Botanists, concerned as they are on the one hand with the swelling pressure of gels, such as cellulose, starch, and agar, and on the other hand with the turgor of the cell vacuole, regard the two processes as distinct because the respective systems in which they take place are mechanically different.

---

[1] Bolam, *The Donnan Equilibria* (London, 1932).

[2] Northrop, *J. Gen. Physiol.*, 1927, 11, 43.

[3] Loeb, *Proteins and the Theory of Colloidal Behaviour* (New York, 1922).

[4] Freundlich, *Colloid and Capillary Chemistry* (London, 1926).

There is a further difficulty in the application of the Donnan equilibrium. The Gibbs' thermodynamic equation requires that at equilibrium the products of the diffusible ions on the two sides of the membrane shall be equal; but osmotic equilibrium demands that the sums shall be equal. Both cannot be true, so that at equilibrium neither condition is fully satisfied; instead, there is a balance between the two forces, neither products nor sums being equal.

The following experiment is an interesting one in which osmotic forces are pitted against imbibition ones: if a series of cubes of, say, 25 % gelatin, all of one size, are placed in sugar solutions of different strengths, water will pass from the gel to the solution, or from the solution into the gel, according to the concentration of the surrounding sugar. One cube in the series may maintain its original size; the others become larger because of the greater imbibition pressure of the gel, or smaller because of the greater osmotic pressure of the surrounding solutions.

There is no doubt but that the Donnan membrane equilibrium applies to living systems. Donnan himself called attention to its probable application to the permeability of the red blood cell. But this does not alter the fact that the osmotic pressure of sugar solutions is a great force in tissues, particularly in plants, and this pressure may readily mask the Gibbs-Donnan effect.

Whatever the final answer to the osmosis *vs.* imbibition controversy may be, it is yet true that the swelling of a gel involves a mechanism, if not a force, which is quite different from that of an osmotic sac. The problem may be stated as follows: whereas the diffusion of water into a jelly against an imbibition pressure may involve the same forces as the diffusion of water into an osmotic sac against turgor, the mechanism which holds the jelly together, preventing solubility, is unique and peculiar to organic colloidal systems capable of exhibiting imbibition.

The problem appears in many forms. There is the extraordinary amount of water held by jellies of less than 1 or 2 % concentration. Gels of germanium or cadmium with but 0·2 % solid matter are quite firm. The amount of water in a jelly-fish is said to be more than 98 %. The holding of so large an amount of water by so small a percentage of solid matter is possible only if the firm gel is built of tenuous, possibly crystalline, fibres which contact each other at specific points. If the molecular fibres are haphazardly arranged we should have a brush-heap: if the arrangement is orderly there will be a three-dimensional lattice or grill.

There are numerous phenomena closely allied to that of the extraordinarily high water content of jellies; such are thixotropy and dilatancy, both described by Freundlich,[5, 6, 7] and much discussed by others.[8, 9] If to a ferric oxide sol a few milligrams of sodium chloride are added, the sol will set to a gel in short time. This gel may be broken down by mechanical agitation, by shaking or stirring, and it then has the remarkable capacity to regelatinise into a gel as firm as the original one; the process may be repeated any number of times.

Thixotropic phenomena occur in protoplasm. The mere stirring of highly viscous protoplasm will often reduce its consistency. The elaborate mitotic figure of a dividing egg, with astral rays and spindle, maintained by structural properties of the protoplasm, may, when subjected to pressure, collapse with great suddenness and reduce the egg to an undifferentiated mass with no vestige of the former mitotic parts; mechanical

---

[5] Freundlich, *The Structure of Protoplasm*, a monograph ed. by W. Seifriz (Ames, Iowa, 1942), p. 85.

[6] Freundlich, *Thixotropy* (Paris, 1935).

[7] Freundlich, *Kolloid Z.*, 1928, 46, 289.

[8] Moyer, ref. 5, p. 23.

[9] Seifriz, ref. 5, p. 245.

disturbance has caused instantaneous and complete thixotropic collapse of the structural arrangement of the cell material.

A further interesting application of thixotropy to biology is that involving anæsthesia.[9] If the culture of a slime mould, such as *Physarum polycephalum*, is placed in a chamber and subjected to suitable gaseous anæsthetic agents, the streaming protoplasm comes to a sudden stop. Recovery takes place within a minute or two, without observable injury. Three anæsthetic agents accomplish this temporary suppression of protoplasmic activity perfectly : carbon dioxide, cyclopropane, and chloroform.

That feature of the anæsthesia of slime mould protoplasm which is of chief interest to us at the moment is the instantaneous cessation of streaming. There is little warning, occasionally a momentary tremor, and then a sudden and complete stoppage of flow with a resumption of normal activity on removal of the anæsthetic agent.

It is only necessary to observe the sudden cessation of protoplasmic movement in a slime mould due to the application of a suitable anæsthetic agent to be convinced that the protoplasm has undergone a pronounced change in its physical as well as in its physiological state, yet no change is evident other than a stopping of flow. In order to ascertain if the protoplasm had, as believed, undergone gelatinisation, pressure was applied by an ingenious technique developed by Kamiya.[10] When air pressure is externally applied to normal protoplasm it may be readily moved, whereas protoplasm under CO_2 anæsthesia cannot be made to flow by pressure. Obviously, then, under CO_2 protoplasm is firm ; when normal it is fluid.

The foregoing facts form the basis of a theory of anæsthesia. The theory states that anæsthesia is due to the thixotropic setting of protoplasm. In the case of higher vertebrates only nerves would be affected ; in slime moulds all the protoplasm sets thixotropically when anæsthetised.

It has been my purpose to show that the structural features which permit the swelling of protoplasm to the point of holding 98 % water are the same as those which permit protoplasm to gelatinise instantaneously when under anæsthesia, namely, by the interlocking of molecular fibres.

Another important phase of the general problem of how water is held by gels is the question of the miscibility or immiscibility of protoplasm in water. Here we have had a controversy of long standing. All agree that protoplasm takes up water with avidity ; but whether it does so osmotically, in the sense of going into solution, or by imbibition, in the sense of soaking up water as does a jelly or a blotter, is the problem. As in the case of Loeb's interpretation of osmosis and imbibition, there may be no difference, but I think there is, and the difference is a very fundamental one ; it is the difference between a sponge and a lump of sugar ; the former takes up water yet holds together, the latter disintegrates. The sugar and not the sponge is miscible in water. Imbibition and solution both depend on the preferential adsorption of water by molecules, but the mechanisms of the two phenomena differ. The molecules of a soluble substance are free when in solution ; those of a substance which imbibes water are held together. Protoplasm, like the sponge, possesses structural continuity. There is no limit to the amount of water in which a miscible substance such as sugar will diffuse ; but a substance such as gelatin which imbibes water and is immiscible in it, takes up a definite maximum amount and then stops. Structural continuity holds it together. This is the case with protoplasm ; it imbibes water but does not go into solution with it.

Dujardin, one of the earliest workers on protoplasm, described protoplasm as a living jelly, insoluble in water. Hertwig referred to it as a viscous, water-immiscible substance ; and Butschli said the same, that protoplasm is not miscible in water. There are, however, supporters of the opposite opinion. They base their view on the rapid diffusion of water when injected into protoplasm. This would also be true if water were injected into a moist but not saturated sponge or blotter. There is much evidence to show that the fibrous structural units of protoplasm form a continuous and rigid but labile framework. This framework, like that of gelatin, exhibits great affinity for water but is not miscible in it. Much of the inorganic and organic matter which permeates the protoplasmic framework is in true aqueous solution.

---

[10] Kamiya, *Science*, 1940, 92, 462.

We come to the crux of our problem, the nature of the framework which gives to protoplasm its structural continuity, and permits swelling, shrinking, flowing, thixotropic setting, stretching, and contraction of the protoplasmic jelly, yet maintains that continuity in structure which is one of the basic prerequisites of living matter.

The earlier concepts on protoplasmic structure, involving granules, alveoli, emulsion globules, and vacuoles,[11] have all been discarded in favour of a reticulum or framework built of linear structural units. Not wholly discarded but greatly restricted and modified is Naegeli's micellar hypothesis involving molecular aggregates fashioned into a symmetrical whole somewhat like bricks in a wall.[12] This was, and still is by some, assumed to be the structure of cellulose.[13] The concept had one great fault, there was nothing to account for a strong bond between the micellæ. A cementing material was assumed to be present, but to attribute the tensile strength of cellulose to an intermicellar substance rather than to the micellæ themselves was poor reasoning. Again the fallacy lay in attempting to build an elastic substance of high tensile strength out of granules. Protoplasm hangs together, and so does cloth; the latter is made of threads. In order to satisfy the elastic and tensile qualities of jellies, the brush-heap theory of interlacing fibres was formulated [11, 12, 14] and out of it arose the concept of symmetrically arranged overlapping fibres. High tensile strength was particularly well explained by such an arrangement.

We come to this conclusion: structurally interpreted, the elasticity of organic material requires that jellies be built of fibres. Low tensile strength, as in Cellophane, suggests a scattered distribution; high tensile strength, as in ramie fibre, requires overlapping. Optical properties, however, indicate discontinuity. Frey-Wyssling [13] shows this to be true for natural cellulose when viewed through crossed Nicol prisms. Also, the examination of natural cellulose with dark-field illumination reveals the presence of short rods, 1·5 to 2μ in length, regularly spaced in parallel rows.[12]

The best interpretation, so it seems to me, of the foregoing somewhat contradictory requirements is that of symmetrically arranged, overlapping fibres concentrated at regular intervals; or, as Meyer [15] and Mark [16] would put it, the points of contact between fibres are not points but regions, there are no free chains, the chains are hydrated and joined over much of their surface. In short, a chain, or part of it, is either linked to other chains or to water. Free chains do not occur owing to the high van der Waals' forces of the —OH or —CONH groups.

Protoplasmic swelling and shrinking involve protoplasmic extensibility and contractility, and all four phenomena involve the adjustment of a structure which is capable of extending and contracting without loss of continuity between fibrous units.

The elasticity of jellies and the high water-holding capacity of thixotropic gels has led to the general acceptance of a linear structural unit in such systems. Contractile properties led to another feature of the structural unit of elastic organic substances, namely, molecular folding. With it our picture is relatively complete.

---

[11] Seifriz, *Protoplasm* (New York, 1936).
[12] Seifriz, ref. 4, p. 1.
[13] Frey-Wyssling, *Protoplasma Monographs*, Vol. 15 (Berlin, 1938).
[14] Seifriz, *Science*, 1938, 88, 21.
[15] Meyer, ref. 5, p. 267.
[16] Mark, *Physical Chemistry of High Polymeric Systems* (New York, 1940).

Proteins are built of polypeptide chains with side chains joining one linear molecule to another. Two classes of proteins are now recognised, fibrous and non-fibrous or corpuscular ones.[17] There is much evidence to support a close relationship between the two with but a small energy difference, so that "it is difficult to avoid the conclusion that fundamentally all proteins are fibrous."[18]

The modern concept of the molecular pattern of proteins is as follows: proteins, according to Astbury,[19] are of four types: extended β-proteins, extended proteins of the collagen type, folded proteins of the α-type, and folded proteins of the supercontracted type. To this scheme of molecular protein types, Astbury[20] and others[21] add one other important structural feature, that of the grid, formed by side chains joining one protein main chain to another. A parallel alignment of grids results in the formation of a three-dimensional lattice, or crystal. It was a haphazardly assembled three-dimensional pattern which the earlier colloid chemists had in mind when they wrote of brush-heaps; these have now become more orderly in design, but not necessarily always so.

We come now to a problem which has long occupied my attention without satisfactory solution until the physical chemists arrived at an understanding of the mechanism of rubber elasticity. The problem has to do with the distinction between substances which have a high modulus of elasticity but low extensibility as does steel, and substances which have a low modulus but high extensibility as do organic jellies such as gelatin, rubber, and protoplasm. Elasticity in terms of modulus is not as closely correlated with the swelling and shrinking of organic jellies as is elasticity in terms of extensibility; the latter is closely associated with the water-holding capacity of jellies, for the two properties rest upon the same structural features.[11]

I should like further to point out that whereas muscle as myosin possesses rubber-like elasticity, the extensibility of muscle as tissue is not great; yet, on the other hand, one of the outstanding characteristics of living muscle is contractility. Now, ordinarily these two properties are coexistent. There is no fundamental difference between elasticity and contractility when inanimate material, such as rubber, is under consideration. A stress-strain curve will, on compression, extend below the line of zero stress as well as above, on extension. But if we are dealing with living matter of the type of muscle, this may or may not be true. The negative temperature coefficient of rubber is not equivalent to the heat contraction of muscle. The Gough-Joule effect[22] need not, therefore, hold for muscle. If heat is applied to muscle, it will contract, but cold does not cause living muscle to expand. This is true because we are dealing with living matter, a dynamic mechanism with self-regulatory powers, and not a passive piece of material. In muscle, any change in metabolism sets up a new state as a result of which tension is developed. The tension thus produced is the result of a new situation, literally a new system, an active state which was not formerly present.

Notwithstanding the foregoing distinction between a living dynamic machine such as muscle in which, on contraction, a new system and not simply a new equilibrium arises, and a non-living organic jelly such as rubber with its high extensibility, there is nevertheless a pronounced similarity in the mechanism involved. Astbury [23] states that the contraction of muscle is due to the supercontraction of oriented myosin; this is also the structural basis of the elasticity of proteins in general. We may, therefore, conclude that although the pronounced contractility of muscle is primarily the formation of a dynamic heat machine, elasticity in the sense of extensibility although feeble is none the less present as a property of the protein myosin. The Gough-Joule effect may, therefore, hold but is completely masked by other qualities of muscle functioning as an independent vital machine.

---

[17] Banga, and Szent-Györgyi, *Science*, 1940, **92**, 514-515.
[18] Astbury, *Trans. Faraday Soc.*, 1940, **36**, 233, 878.
[19] Astbury, *Fundamentals of Fibre Structure* (London, 1933).
[20] Astbury, *Tabulae Biologicae*, 1939, **17**, 90.
[21] Meyer, *Die hochpolymeren Verbindungen* (Leipzig, 1940).
[22] Joule, *Trans. Roy. Soc.*, 1859, **149**, 91.

Before summarising my views on the structural qualities of living matter, such as are responsible for great elasticity, swelling and correlated properties, I should like to call attention to a closely related property of rubber, namely, its remarkably great extensibility, reaching a maximum of 900 %, which has long defied analysis. More moderate extensions of proteins have been interpreted by Astbury [23] as due to the supercontraction of oriented linear molecules. Similar suggestions have been made by Mark [24] and Meyer.[21]

I should like to summarise the views expounded by the foregoing physical chemists, possibly in a slightly modified form, but at least in other words. The tendency of rubber-like matter to take its original form after deformation can be ascribed to the fact that the chain molecules are themselves capable of deformation, and that parts of adjacent chains can slide over one another. Stretching and reorientation force molecules, which, when relaxed, may assume any chance position and shape, to take on less possible forms and arrangements. Due to their kinetic energy, the thermodynamic probability that the molecules will assume a disordered state becomes greater as soon as the external force causing deformation disappears. It follows from this that only systems containing chain molecules, the inner mobility of which is not prevented by crystallisation or micelle formation, can possess rubber-like elasticity.

And now we may carry our concepts directly to protoplasm and say that the contractility of protoplasm is due to the supercontraction of its structural proteins through the folding of molecular fibres symmetrically aligned and joined one to the other by side chains so as to form a three-dimensional lattice. As this is a general conclusion of far-reaching significance in biology, it may be restated as follows: contractility wherever it occurs in animate nature, whether in the most elementary form of protoplasm or in highly specialised tissue, is due to the shortening of protein fibres by molecular folding. The energy for this work is supplied by the oxidative processes of the cell.

As I have stated, the structural features responsible for extensibility and contractility are likewise the mechanism which permits swelling and shrinking. The biological and medical applications of protoplasmic swelling and shrinking and the molecular mechanism involved are many. Aside from the non-Newtonian viscosity of protoplasm, its tensile strength, elasticity, contractility, swelling and shrinking, there are a number of physiological properties which in varying degree involve the same structural features; these include growth, sweating when it is a case of syneresis, ageing, body form, and rhythmic pulsations. Muscular activity, now generally regarded as due to the contractility of molecular fibres, has been viewed as a case of swelling and shrinking. Hofmeister believed that the delicate protein fibrils in the muscle of a bee's wing imbibe water and give it up again, by syneresis, rapidly enough to permit contraction and relaxation two hundred times a second. Mirsky [25] has correlated muscular contraction with dehydration. To-day we may still state that the older hypothesis based on swelling is still true, and the modern hypothesis based on molecular contraction likewise true; the one, hydration, is the chemical means by which the other, the physical change of contraction, is brought about.

---

[23] Astbury, *Trans. Faraday Soc.*, 1940, 34, 871.

[24] Mark, *Kautschuk*, 1930, 6, 2.

[25] Mirsky, *Science*, 1936, 84, 333.

Water content has long been known to be correlated with ageing, but whether we get old because we dry up or dry up because we get old is a question. In any case, it appears that the capacity of our protoplasm to hold water decreases with age. Ageing may be due to a less hydrated condition of the protoplasm, and to the gradual going over of albumin into globulin, which albumin does spontaneously.

Body form may, in certain instances, be due to shrinking as a result of dehydration. Hatschek has suggested that the gastrulation of an embryo takes place in the same manner and for the same reason as does the invagination of a two-thirds segment of a sphere of gelatin on drying.

Rhythm in nature, particularly the rhythmic contractility of protoplasm, has received much of my attention of late.[26, 9] I have expressed the belief that it is the mechanism responsible for protoplasmic flow; more than this, it is a fundamental property of all protoplasm, retained in mature tissues where it is functionally necessary, as in heart muscle. Rhythmic contractility is best known in the tissues of higher organisms, but it is a basic quality of the most primitive forms of life as well. It has been possible to measure the pressures of forces involved in the rhythmic contraction and expansion of the protoplasm of the *Myxomycetes*.[10] Kamiya [27] was able not only to measure pressures but to analyse the irregularities in the graphs of protoplasmic rhythms in the same way that the physicist analyses beats in harmonics. As a result of such studies in protoplasmic harmonics Kamiya concluded that the protoplasm of a slime mould is a polyrhythmic system.

One of the most satisfying features of recent research on protoplasm is the complete agreement among, and perfect correlation between all properties, physical and chemical, studied. If protoplasm is found to be doubly refractive, one can be fairly certain that it will also prove to be non-Newtonian in character, i.e. show variable viscosity. And in turn, anomalous viscosity is almost certain to mean that the material is elastic. Elasticity, contractility, double refraction, tensile strength, thixotropy, the coexistence of rigidity and fluidity, owe their existence to linear structural units. And again we find correlation between the foregoing properties and water content. Scarth [28] says that birefringence in protoplasm tends to disappear with a high degree of hydration and reappear on dehydration. Mirsky [25] has correlated fertilisation with dehydration.

No property of organic matter, especially living matter, is more important than the constant shifting of bonds between fibrous structural units. It is this quality in a dynamic system such as protoplasm which gives to it both fluidity and structural continuity. In any vital chemical cycle, such as that in muscle during which accessory molecules go in and out of combination with the side chains of the protein main chains, the molecules take up a cycle of configurations. The hydrogen bond makes this easy shifting of ties possible.[29] Astbury [23] ascribes to the hydrogen bond the important function of protein hydration and therefore protein tautomerism.

The paramount vital importance of the swelling and shrinking of tissues may be further emphasised by pointing out some other roles which these processes play in normal and pathological physiology. There are: the hydrophilic swelling of the parenchymatous tissue of liver and kidney, the shrinking of muscle in poliomyelitis, the atrophy of muscle from disuse when a limb is immobilised in a cast, hypertrophy of over-use in heart and striated muscle due to the swelling of cells, œdema where there

---

[26] Seifriz, *Science*, 1937, 86, 397.

[27] Kamiya, ref. 4, p. 199.

[28] Scarth, ref. 4, p. 99.

[29] Pauling, *The Nature of the Chemical Bond* (Ithaca, New York, 1940).

is an accumulation of fluid in cells as well as in interstitial spaces, the swelling and shrinking of active glands illustrated by the imbibition of water and the secretion of fluid by the alveoli of salivary glands, shrinkage of tissues in pituitary cachexia, and finally, the most fundamental of all forms of water uptake and water loss, namely, that by protoplasm in its precise control of water content.

University of Pennsylvania,
Botanical Laboratory,
38th St., Woodland Ave.,
Philadelphia 4, Penna., U.S.A.

SWELLING AND SHRINKING OF MAMMALIAN RED CELLS: DEPARTURES FROM IDEAL OSMOTIC BEHAVIOUR.

By Eric Ponder.

Received August 12th, 1946.

1. Volume Changes.

If we start with a simple conception of the structure of the mammalian red cell and suppose that it is a balloon-like body surrounded by a semi-permeable envelope or membrane which does not resist the changes in volume brought about by osmotic forces, that its interior is initially in osmotic equilibrium with the surrounding plasma, and that the volume changes which occur in a hypertonic or hypotonic medium are the result of the transfer of water alone, we can expect that the new volume V which the cell attains in a very large volume of a medium of tonicity T will be

$$V = W(1/T - 1) + 100 \quad . \quad . \quad . \quad . \quad (1)$$

where the initial volume of the cell in an isotonic medium is denoted by 100, the percentage of cell water by W, and where the tonicity T is defined as the ratio of the depression of freezing point of the suspension medium to that of plasma itself. Expression (1) is one form of the " van't Hoff-Mariotte law " for ideal osmotic swelling, and if the behaviour of the cell can be described in this simple way we speak of it as a " perfect osmometer " ; the following relation between tonicity and volume, for example, would be found for a perfect osmometer with $W = 70$:

T . . .	1·5	1·4	1·2	1·0	0·8	0·6	0·5
V . . .	76	80	88	100	116	146	170

There has been much controversy as to whether the mammalian erythrocyte behaves in this way when immersed in hypotonic or hypertonic media, and the divergence of opinion has arisen on a variety of grounds. There have been errors and uncertainties in the methods for measuring red cell volume, and it is only recently that the conditions have been defined under which the results of different methods can be expected to agree with each other.[1] Further, it has not been sufficiently appreciated that the osmotic behaviour of the red cell, as well as some of its other permeability properties, depends on the nature of the environment to which it is exposed. Loss of K and of other osmotically active substances is, for example, an almost constant prolytic phenomenon,[2] and the osmotic behaviour of red cells in hypotonic NaCl is not the same as their behaviour in hypotonic plasma.[3] Lastly, the deviations from the expected osmotic behaviour have not been sufficiently seriously considered even when they have been observed to occur. Rather than conclude that simple relations do not exist, dubious subsidiary hypotheses and corrections have been introduced to bring the volumes observed in hypertonic and hypotonic media into agreement with the volumes expected, and so to permit the retention of the conception that swelling and skrinking occur as a result of the transfer of water alone.

---

[1] Ponder, *Medical Physics* (The Year Book Publishers, Chicago, 1944), p. 1203.
[2] Davson and Danielli, *Biochem. J.*, 1938, 32, 991.
[3] Ponder and Robinson, *J. Physiol.*, 1934, 83, 34.

When the results which have been obtained are collected[4] it becomes apparent that the mammalian red cell sometimes behaves as a very good, if not perfect, osmometer, but that it sometimes swells and shrinks to a smaller extent, so that the volume V is given by

$$V = RW\,(1/T - 1) + 100 \qquad . \quad . \quad . \quad (2)^{5}$$

where R is a constant which is introduced so as to reconcile observation with simple osmotic theory. For $R = 0{\cdot}5$, for example, the relation between V and T for a cell with $W = 70$ would be

T . . .	1·5	1·4	1·2	1·0	0·8	0·6	0·5
V . . .	88	90	94	100	108	123	135

It has been found that rabbit red cells from heparinised[6] or defibrinated[3] blood behave as almost perfect osmometers ($R = 0{\cdot}9$ to $1{\cdot}0$) in hypotonic plasma from heparinised blood or in hypotonic serum from defibrinated blood, but that cells from oxalated blood behave as imperfect osmometers with $R = 0{\cdot}5$ to $0{\cdot}7$ in hypotonic oxalated plasma. Cells in hypotonic saline media also tend to behave as imperfect osmometers,[3, 4] and in general the value of R depends on the conditions under which the volume measurements are made. Moreover, it is found that the value frequently depends on the degree of hypertonicity or hypotonicity, so that a red cell which behaves as a perfect osmometer when T is near unity may be a poor osmometer when T is 1·5 or 0·5.

While there is no question as to the observations, there has been a great deal of discussion as to the meaning of R. One explanation is that R measures the fraction of the cell water which is free, the remainder being bound. This idea was put forward at one time to account for the anomalously small volume changes in muscle, but was shown to be untenable by Hill.[7] To account for the values of R obtained in experiments on the swelling and shrinking of red cells, this explanation would require that the quantity of free water should undergo variations from about 50 % of the total cell water to about 100 % of it. In the red cell from defibrinated blood at least 95 % of the water is free in the sense that it is solvent water,[8] and it seems unlikely at first sight that the necessary variation in the free water would be brought about by merely oxalating the blood ; discussion of this point is best postponed, however, because the term " free water " may have a special meaning (see below).

A second explanation for the appearance of the constant R is that the red cell does not come into equilibrium with its surroundings by the exchange of water alone, but also by the loss, or gain, of osmotically active substances, e.g. of K in the case of the red cells of the rabbit or of man. If equilibrium were reached by the transfer of water together with a loss or gain of K, the quantity of the ion lost into a very large volume of

---

[4] Ponder, *Cold Spring Harbor Symposia*, 1940, 8, 133.

[5] If the volume of the suspension medium is limited, the expression becomes

$$V = RW\,.\,\frac{p - pT}{pT + 1} + 100 \qquad . \quad . \quad . \quad . \quad (3)$$

in which p is the ratio of the volume of the suspension medium to the volume of the cell water W.

[6] Guest and Wing, *J. Clin. Invest.*, 1942, 21, 257.

[7] Hill, *Proc. Roy. Soc. B.*, 1930, **106**, 477.

[8] Macleod and Ponder, *J. Physiol.*, 1935, **86**, 147.

a hypotonic medium, or gained from a very large volume of a hypertonic medium, would be $100\ (1 - R)(1 - T)\ \%$ of the total K initially contained in the cells; for $R = 1{\cdot}0$, there is no loss, and swelling and shrinking take place as a result of water exchange alone, while for smaller values of R there should be increasing losses to a hypotonic medium or gains from a hypertonic medium. This hypothesis is not satisfactory in a quantitative sense, for the amounts of K lost into hypotonic media are not large enough to account for the small values of R observed.[3, 4, 9, 10] Nevertheless, losses have been observed,[11] and that their occurrence has a place in the complete description of the processes underlying the volume changes is all the more certain when we consider that the cells are no longer regarded as impermeable to cations,[12, 13] and that the loss of K occurs frequently as a preliminary to lysis.[2]

A third explanation is that the anomalous osmotic behaviour is due to the osmotic activity-concentration curve of hæmoglobin being convex to the concentration axis even when the concentrations are expressed in terms of the water which is solvent water for crystalloids.[14] Considerations such as this illustrate the fallacy of expecting that simple osmotic relations necessarily apply to the concentrated systems in the red cell interior.

A still more important factor is the state of the red cell interior as shown by the presence or absence of a coarse type of crenation. This variety of crenation (crenation of type II) occurs in the red cells of blood which has been allowed to stand, in oxalated blood, and under conditions in which the hæmoglobin and perhaps other components of the cell tend to become gelated or paracrystalline. The change from the dispersed to the paracrystalline state confers a birefringence on the crenated cells.[15] Crenation of type II should not be confused with the fine crenation (crenation of type I) which takes place during the shape transformation of red cells from discs to spheres as a result of the increase in *p*H which occurs between a glass slide and coverglass,[16, 17, 18] or as a result of the action of lysins or of sphering agents such as lecithin.[19, 20] Nor must crenation be taken as evidence of a diminution in red cell volume; we rather regard it as the result of a loss of the ability of the cell to maintain its special shape. The coarse form of crenation, however, is associated with anomalously small volume increases in hypotonic media, so that there is a relation between the extent of crenation, as observed microscopically, and the value of R.[21] Crenated cells give a uniformly lower R-value, the swelling of crenated cells from oxalated blood in hypotonic oxalated plasma being described by expression (2) with $R = 0{\cdot}5$ to $0{\cdot}7$, while the swelling of the discoidal, uncrenated cells from heparinised blood in hypotonic heparinised plasma is more nearly that of a "perfect osmometer," with R-values between $0{\cdot}9$ and $1{\cdot}0$.

The view that the variations observed in osmotic swelling are essentially

[9] Davson, *Biochem. J.*, 1934, **28**, 676.

[10] Davson, *J. Cell. Comp. Physiol.*, 1937, **10**, 247.

[11] The losses of K given by Ponder and Robinson, *Biochem. J.*, 1934, **28**, 1940, are considerably larger than those given by me in 1940. Davson's values are intermediate. There is a good deal of variation in the results of superficially similar experiments on the loss of K from red cells into hypotonic media, and the question needs further investigation.

[12] Harris, *J. Biol. Chem.*, 1941, **141**, 579.

[13] Dean, Noonan, Haege and Fenn, *J. gen. Physiol.*, 1941, **24**, 353.

[14] Roepke, and Baldes, *J. Cell. Comp. Physiol.*, 1942, **20**, 71.

[15] Teitel-Bernard, *Arch. Roum. Path.*, 1932, **5**, 389.

[16] Ponder, *J. Exp. Biol.*, 1929, **6**, 387.

[17] Furchgott, *ibid.*, 1940, **17**, 30.

[18] Furchgott and Ponder, *ibid.*, 1940, **17**, 117.

[19] Ponder, *ibid.*, 1936, **13**, 289.

[20] Ponder, *ibid.*, 1943, **19**, 220.

[21] Ponder, *J. gen. Physiol.*, 1944, **27**, 273.

due to changes in the state of the red cell interior derives strong support from the description of an extreme case of anomalous osmotic behaviour associated with the gelation of rat red cells in isotonic sodium citrate at low temperatures (4° to 9° c.). The resistance of these cells to hypotonic hæmolysis increases so much that after several days very little swelling (R = 0·05 to 0·15) is observed in hypotonic media, and most of the cells do not hæmolyse even in cold distilled water.[22] The failure to swell is not due to a loss of K, for the concentration inside is at least 20 times that outside; we have therefore to conclude that the condition in the interior is one in which the activity of K approaches zero, as it would if the system were crystalline or paracrystalline. These paracrystalline red cells do not undergo the usual shape change from disc to sphere when lysins are added; they seem to be too rigid, and after they have lost their hæmoglobin a substantial granular debris remains. The paracrystalline cells are also strongly birefringent, but on warming the birefringence disappears except at foci along the rim, to reappear throughout the entire cell on cooling to 4° c. The osmotic properties, on the other hand, appear on warming to 37° c. and disappear again within a few minutes on cooling; the paracrystalline state, with its birefringence, rigidity, and almost complete inability to undergo osmotic swelling, is accordingly a reversible one.

These observations show that the osmotic behaviour of the red cell depends largely on the conditions in its interior. It is usual to approach the problem of the swelling and shrinking of the cell from the standpoint of what would occur in dilute solutions separated by a semipermeable membrane, and in this way to develop expressions such as (1) on the basis of a simplified osmotic theory. The interior of the normal discoidal cell, however, can scarcely be considered as a dilute solution, for the hæmoglobin molecules are closely packed together with an average distance of only about 8 A. between them.[22] Simplified osmotic laws are not likely to apply to such a system, which can be quite properly regarded as an expanded crystal. The tendency for the system to undergo gelation will also be apparent, and since the gelation is reversible, the normal discoidal red cell and the paracrystalline red cell can be looked on as two metastable forms, the orderliness of the interior being greater in the paracrystalline form than in the normal disc, but greater in the normal disc than in a hæmoglobin solution. In the paracrystalline form, the activity of K in the interior would approach zero, while in the interior of the normal disc it would approach its value in a 34 g./100 ml. hæmoglobin solution. Between these extremes there might be a continuous series of values corresponding to a series of almost completely reversible states of the interior. These values for the activity of K could be translated into R-values, and so we would have a series of values depending on the extent to which the paracrystalline state had developed in the individual cell and the fraction of the total number of cells involved in the change. As all values of R between 1·0 and zero have now been observed experimentally, it seems that the osmotic properties characteristic of one state can pass continuously into those of the next.

The conclusion that the osmotic behaviour is a function of the state existing in the red cell interior should not make us lose sight of the fact that under special circumstances the cell behaves as a "perfect osmometer" with an R-value approaching unity.[4, 8, 21] This is the characteristic behaviour of cells from heparinised or defibrinated blood when suspended in plasma of various tonicities, but it is a behaviour which ought to be demonstrated experimentally in any given system, and not simply assumed. The fact that swelling and shrinking can be described by expressions of the type of (1) and (2) should not, moreover, be relied upon as an argument for the volume changes being due to the osmotic transfer of water across a semipermeable membrane; the range over which the volume changes can be observed is very small ($T = 2{\cdot}0$ to $T = 0{\cdot}5$), and the swelling and shrinking might be described by equations set up on the basis of osmotic transfer even if the underlying process were essentially an imbibition process, such as occurs in gelatin gels.

---

[22] Ponder, *Proc. N.Y. Acad. Sci.*, 1946 (in press).

2. Shape Changes.

The assumptions upon which expression (1) is set up do not contain any allowance for the structural components of the red cell offering a resistance to osmotic forces. The shape changes which individual cells undergo in hypotonic systems, however, cannot be accounted for unless one recognises their rigidity of form.

In hypotonic plasma or saline, neighbouring cells can be observed to undergo swelling to different extents and in different ways. A swollen cell may lie side by side with a crenated one, and some cells retain their biconcave shape whereas others become spheroidal or cup-shaped. Presumably the different shapes observed during swelling correspond to different rigidities of cell structure and to different degrees of resistance offered to the osmotic forces. Generally speaking, however, the area of the cell increases only very slightly as swelling proceeds, and the increasing volume is accommodated by a shape change in which the length/breadth ratio becomes less. A point is then reached at which some structural component seems to give way and at which the cell becomes a sphere with a considerable reduction in surface area ; this yield point corresponds to a smaller volume increase in the case of some cells than in the case of others. With further increase in volume, the area of the sphere increases until the cell hæmolyses at its " critical volume " ; when this happens, the osmotic forces disappear, and the hæmolysed ghost returns within less than a second to a shape more or less the same as that of the biconcave disc from which it was derived. This sequence of events is incompatible with anything but a body which has rigidity of form.[21]

In the case of red cells which have undergone gelation, an approximate relation between the value of *R* and the resistance to swelling offered by structural components can be established by supposing that the crenated cell is a uniform gel and calculating the bulk modulus which it would have to possess in order to account for the small osmotic swelling. The values for the bulk modulus work out at about the same as those for gelatin gels. It should be emphasised, however, that this is only one way of accounting for the anomalously small swelling. An equally good way is to suppose that the change to the paracrystalline state involves a decrease in the activity of the osmotically active substances in the red cell interior, or even in the amount of " bound water," the term " bound water " now having the special meaning of " water rendered incapable of taking part normally in osmotic processes because of the restraints imposed by the paracrystalline structure." There is no satisfactory way, in the meantime, of distinguishing between these possibilities, but it should be noticed that they all account for the departure from ideal osmotic behaviour in terms of the limitations imposed by red cell structure.[22, 23]

3. Red Cell Fragility and its Relation to Osmotic Swelling.

Changes in the fragility of red cells are closely associated with changes in their osmotic behaviour, and although the fragility is often spoken of as if it were a simple property, it depends in a complex way on a number

---

[23] In experiments with red cells in heparinised or defibrinated plasma, the *R*-values tend to be more nearly 0·95 than 1·0. The difference is not likely to be due to experimental error, and it is questionable what it represents. Some of it may be accounted for by the binding of a small amount of water to hæmoglobin ; the remainder of the difference, if any remains to be accounted for, would represent a structural resistance to the forces which produce swelling.

of variables. To determine it, red cells are placed in a series of relatively large volumes of media of decreasing tonicities, separated by intervals of about $T = 0{\cdot}02$, and the tonicity in which there is just commencing hæmolysis determines the fragility. Some investigators prefer to use the tonicity in which there is complete hæmolysis and some the tonicity in which there is 50 % hæmolysis; these differences of procedure are not important so far as this discussion is concerned.

It will be clear from expressions (1) and (2) that the tonicity T in which a given amount of lysis occurs depends on the value for W, the value of R and the value of V at which lysis occurs, or the "critical volume" for the cell under consideration. This last value approximates to the volume of the sphere which has the same surface area as that of the discoidal red cell before swelling begins.[24, 25] The average discoidal cell, for example, has an initial volume of 60 μ^3 in isotonic plasma, and its surface area, 105 μ^2, is considerably in excess of 75 μ^2, the minimum surface for the enclosed volume. When the cell undergoes swelling in a hypotonic medium, it increases in volume and changes shape, becoming more spherical until it is finally a sphere with a surface area of 105 μ^2, i.e. with a surface area the same as that of the original discoidal form. The volume of such a sphere is 100 μ^3, and the increase in volume of 167 % is possible without stretching of the cell surface. Looking at the matter from a simplified standpoint, the cell hæmolyses when this volume is reached, and the volume 100 μ^3 or 167 % of the initial volume V_0 is the "critical volume" V_c. The swelling preceding hæmolysis accordingly involves a rearrangement of the components of the red cell, but not an appreciable extension of its surface layers.

If the maximum area for a swollen sphere with volume V_c is the same as the area of the discoidal form of initial volume V_0, the extent to which swelling can occur without lysis will be determined by the extent to which the area of the discoidal form exceeds the minimum area for the enclosed volume V_0. The extent to which swelling can occur without stretching the surface and hæmolysis is accordingly a function of the initial shape of the cell, cells which are more nearly spherical being capable of less swelling than cells which are flatter. The larger the critical volume, the lower the value of the critical tonicity T_c at which this volume occurs, and so the fragility of the cell, measured in terms of T, has come to be looked upon as primarily dependent on red cell shape.[26, 27] It is true, generally speaking, that red cells which are more nearly spherical (spherocytes) have low fragilities and hæmolyse at comparatively high values of T, whereas flat cells (platycytes) hæmolyse at lower tonicities.

Several factors operate to modify this simplified description. The first is that the swollen volume enclosed by the same surface as that of the original disc is not always reached before the cell hæmolyses. Red cells in hypotonic saline hæmolyse at a smaller critical volume than do the red cells of the same animal in hypotonic plasma,[3] and this one instance, even though it is the only one yet described, is sufficient to put one on one's guard against assuming that the critical volume is independent of the nature of the hypotonic medium in which the swelling takes place.

The influence of the remaining factors will be appreciated if we rearrange expression (2) as

$$T_c = \frac{RW}{100(V_c/V_0 - 1) + RW} \quad . \quad . \quad . \quad (4)$$

The critical tonicity T_c in which lysis occurs depends on the value of V_c/V_0, a complex function of cell shape, but also on the values of W and

---

[24] Castle and Daland, *Arch. Int. Med.*, 1937, **60**, 949.
[25] Ponder, *J. Exp. Biol.*, 1937, **14**, 267.
[26] Gänsslen, *Deutsch. Arch. klin. Med.*, 1922, **140**, 210.
[27] Haden, *Amer. J. Med. Sci.*, 1934, **188**, 441.

of R. When we are comparing the osmotic behaviour of red cells in different hypotonic media, it cannot be assumed that R is constant, and when we are comparing the behaviour of red cells from different animals or from persons under different conditions, it cannot be assumed that either R or W is constant. The values of R obtained for normal human red cells from heparinised blood, for example, tend to be from 0·9 to 1·0, whereas the values found for cells from cases of sicklemia and Mediterranean anæmia are much smaller; these variations are reflected in the value for the critical tonicity just as clearly as variations in red cell shape would be. Many similar instances occur, and still more complicated situations, in which R is a function of T itself, have been described. Since the fragility depends on so many variables, the interpretation of its variations as unequivocal evidence of variations in red cell shape is apt to be misleading.

The Nassau Hospital,
Mineola, N.Y.

*GENERAL DISCUSSION**

Mr. N. Goldenberg (*London*) (*communicated*): The picture of the swelling of starch granules given by Dr. Badenhuizen does not include a description of the fundamental mechanism of swelling, which takes place in the way one would expect for a highly-branched polymer, and which appears to hold whether the starch granule has a layer structure or not. Thus, in the unbroken granules, the amylose and amylopectin macromolecules are held together, presumably by means of hydrogen bonds between neighbouring hydroxyl-groups, in a structure that is similar to that obtaining in cellulose and is made up of crystalline and amorphous regions. A certain energy of activation is required to break up the cohesive forces thus set up; this can be achieved either by heat or by the action of suitable chemicals. Below a certain temperature, the molecular framework does not break up and consequently the granules retain their microscopical appearance, birefringence and X-ray diffraction pattern and show a limited reversible swelling. Above this temperature, the network is broken, the many hydroxyl-groups previously tied to one another are free to combine with water molecules, the starch macro-molecules become heavily hydrated, the granule swells considerably and the familiar starch paste results. A similar picture is obtained when suitable chemicals are used.

Further, it is not certain whether it is justifiable to accept a layer structure in the original untreated granules when relatively drastic treatments have been used to demonstrate its presence, since the layer structure may have been produced by these treatments.

Dr. W. W. Barkas (*Princes Risborough*) said: While agreeing with Seifriz that water control in protoplasm is probably not just a simple case of osmosis, I would question whether there is really any fundamental difference between osmosis and imbibition even if Freundlich did say it was false to assume their identity. The laws of osmosis can be deduced thermodynamically from data on vapour pressure lowering and volume changes by methods which are independent of any assumptions regarding the mechanism involved, i.e. whether the water is " going into solution " or being " soaked up as in a sponge." These are two sub-divisions of the osmotic process and a third very important one occurs where the solute does " go into solution " but is prevented from dispersing by the fact that the molecules of the solute possess, as the author says, structural continuity. This is no less a solution, though not by any means an ideal one, in the thermodynamic sense. In wood it is possible to break down the structural continuity of the unit cell and the cell wall material then disperses into a much larger volume, though complete dispersion may be prevented at this stage also by the fibrillar sub-structure of the material.

---

* On four preceding papers.

[1] Stoves (in press).

Mr. L. Peters (*Leeds*) said : Seifriz contrasts the Donnan-Gibbs thermodynamic equations of equilibrium which demand that the products of the diffusible ions on the two sides of the membrane shall be equal, with the osmotic conditions of equilibrium that the sums shall be equal.

The exact thermodynamic conditions for equilibrium for solutions of any concentration have already been deduced by Donnan and Guggenheim [2] using the method of Willard Gibbs which requires that the chemical potentials of each species of particles shall be equal on both sides of the membrane. These equations contain terms which represent the activity of the solvent as well as that of the ionic species involved and hence they express the relationship between the osmotic and the ionic equilibria. There is, thus, no opposition between the osmotic and the " Gibbs " thermodynamic equilibria as Seifriz suggests.

The order of magnitude of the effect due to osmotic pressure as compared with that due to an ionic membrane equilibrium can be seen from the case where the dilute solution is molar in concentration with a fully ionised salt and (say) ten molar in the other. Even under these extreme conditions the mole fraction of the water in the concentrated solution is only 16 % less than that of the dilute solution as against a 1000 % difference in the ionic concentrations.[3] Such high concentrations are rarely met with in the biological world and it is, therefore, a very close approximation to the exact treatment to ignore the osmotic effects when only ionic effects are present.

In the discussion, previous speakers mentioned the difficulty of explaining in terms of the Donnan effect the preferential absorption of K^+ ions by red blood cells from solutions containing relatively small concentrations of potassium compared with those of sodium ions.

If the " internal " volume of the cells is very small, the degree of hydration of the Na^+ ion may decrease the mole fraction of free water in the cell so much more than the unhydrated K^+ ion that the equilibrium is only possible by the preferential entry of the latter. But in any case, it is futile to expect *equilibrium* equations to apply to a system in which a process consuming energy is proceeding, e.g. in the living cell. This is shown by the radical redistribution of ions which takes place when the cell ceases to live.

Dr. R. K. Schofield (*Harpenden*) said : An investigation in progress on the negative adsorption of chloride in jute fibres shows promise of giving information about the relative concentration of the cations close to the surfaces to which the COO^- groups are attached. With potassium chloride the negative adsorption is less than with sodium chloride, presumably because the maximum concentration of potassium ions is greater than that of sodium ions. There is no reason to regard this effect as specifically biological : it is doubtless a consequence of ion size.

Professor A. R. Ubbelohde (*Belfast*) said : In discussing the osmotic equilibrium between aqueous solutions inside and outside a living cell, or between an aqueous solution closely enmeshed by molecules of soap and an aqueous solution in bulk, it is important to keep in mind that water is a complex solvent, and contains H_2O molecules in various degrees of association.

For purposes of discussion we may represent the equilibrium between simple and " polymerised " H_2O molecules by formal chemical equations of the type

$$(H_2O)_n \rightleftharpoons nH_2O,$$

where n can have a range of values. The forces causing association are of the weak hydrogen-bond type, which lead to *para*-crystalline tetrahedral arrangements of clusters of H_2O molecules in liquid water, as discussed by Fowler and Bernal. Experimental estimates of the entropy and free energy changes required for changes in the degree of polymerisation can be made, for example, from data on the latent heat of condensation and vapour density of water vapour at various temperatures. The values are such that the position of equilibrium between simple and polymerised H_2O molecules may be shifted to a considerable extent by comparatively small changes in the system.

---

[2] *Z. physikal Chem. A*, 1932, **162**, 346.

[3] Guggenheim (private communication).

The " water " inside a cell, or in the neighbourhood of fatty acid molecules, may be expected to differ markedly in molecular composition from water in bulk. Allowance for this difference in molecular composition must be made in examining osmotic equilibrium; the osmotic pressure of aqueous solutions after they have been squeezed out of cells is not immediately comparable with the osmotic pressure of the same solutions within the cell walls, because the " solvent " may alter in molecular composition when it is squeezed out.

To test this suggestion, direct experimental evidence is required about the degree of polymerisation of H_2O within such enclosures, since any other explanation of the remarkable osmotic effects observed may be superfluous. The chemical analogy between H_2O, NH_3 and H_2S suggests that *in situ* measurements of the dielectric constant and dielectric relaxation of the " water " within the enclosures may give a fairly sensitive measure of the degree of polymerisation. In support of this suggestion it may be noted that the dielectric constants of simple NH_3 and H_2S molecules (1·44; 0·93 D. respectively) and simple H_2O molecules (1·85 D.) are not very different, so that the anomalous properties of liquid water in bulk, such as the high boiling point (H_2O 100° C., NH_3 − 33·4° C., H_2S − 60·7° C.), and high dielectric constant, are in part due to its containing clusters of H_2O molecules linked by hydrogen bonds. Such clusters are apparently not formed in liquid NH_3 or H_2S to any extent.

At a guess, the proportion of H_2O clusters would be expected to fall off in the proximity of the walls of the enclosures, and this should have a measurable effect on the dielectric properties. Changes in the degree of polymerisation would likewise affect the thermodynamic activity of hydrated ions to a different extent from the effect on less hydrated ions, i.e. to a different extent for sodium ions, from that for potassium ions.

Other direct experimental measurements on the " water " inside the enclosures might also be used to verify the degree of polymerisation of the water. The essential point of these remarks is that unless we know how far this " water " differs in molecular composition from water in bulk, it is difficult to evaluate alternative theories of the effects observed.

Dr. D. H. Bangham (*London*) (*communicated*) : Whilst otherwise endorsing Prof. Ubbelohde's remarks about the modification of the state of aggregation of water molecules in close proximity to non-polar surfaces, I would place rather less emphasis upon the possible role of the unassociated molecules. Saturated or nearly saturated water vapour forms on non-polar surfaces (e.g. of charcoal), films which are structurally different from ordinary liquid water and which must offer a different thermodynamic environment to ions and other dissolved species. The water molecules in these films are markedly associated, but it is their geometrical arrangement which appears to differentiate them from ordinary water; they appear to form small closed groups or clusters between which only weak attractive forces are acting [4]

When a non-polar surface is *immersed* in a bulk of water, the aggregational pattern of the molecules at the interface must, it is true, be partly dominated by that of the molecules more remote from the interface, i.e. by that characteristic of the bulk liquid. In a system where the amount of liquid remote from an interface is relatively small, however, this would not be the case and the abnormality of the properties would be accentuated. This would apply, in particular, to the cell contents which, though immersed in water, are separated from it by a membrane or membranes.

---

[4] Bangham, *Nature*, 1946, 157, 733.

Dr. M. F. Perutz (*Cambridge*) said: My observations on hæmoglobin crystals lend further support to the views expressed by Prof. Ubbelohde and Dr. Dervichian. According to Ponder, red cells contain 34 g./100 ml. hæmoglobin which amounts to 26 % by volume. This is only a little less than in hæmoglobin crystals where the hæmoglobin molecules occupy 44 % of the unit cell volume at the maximum degree of swelling. Hæmoglobin crystals do certainly not behave as osmometers, as their volume is independent of the concentration of neutral electrolyte in the medium. On the other hand, only 0·3 g. H_2O/g. protein behaves as bound water, in the sense that it is not available as a solvent. If the amount of bound water in the red cells were the same as in the crystals, the hydrated hæmoglobin molecules would take up 34 % of the cell volume, leaving 66 % still available as solvent water. Nevertheless it would be surprising if this water behaved as a thermodynamically normal solvent, in view of the small distances between the strongly polar surfaces of neighbouring hæmoglobin molecules. It would be interesting to see if the crenated, birefringent cells observed by Ponder showed discontinuous swelling similar to the hæmoglobin crystals.

Prof. J. D. Bernal (*London*) said: I was extremely interested in Prof. Ubbelohde's remarks on the polymerisation of water molecules. It seems to me that there are possibilities of measuring the degree of polymerisation both by infra-red observation, by X-ray scattering and by volume relations. It should also not be impossible in the simpler cases such as those where long chain molecules and hydrophobe groups are present to calculate from models the kind of arrangement of water molecules likely in their immediate neighbourhood. A combined attack by all these methods would, I think, be well worth while. The work of Fowler and myself could now be checked by many more recent experimental determinations, particularly by those of Magat and Finbak.

THE SWELLING OF REINFORCED PLASTICS.

By F. T. Barwell and K. W. Pepper.

Received 1st August, 1946.

Many industrial plastics behave as swelling gels, that is, for every value of the vapour pressure of adsorbate an equilibrium quantity thereof is sorbed and the gel in consequence assumes a definite volume. If any external restraint to swelling is applied, the equilibrium sorption is reduced and a swelling pressure is exerted on the restraining influence. This behaviour is important in engineering applications of plastics, where three conditions may occur.

(*a*) Where no restraint is afforded, changes in dimensions may result. For example, the field of application of water-lubricated plastics bearings is limited by their swelling. Generally, parts may not fit, although initially compatible, because of differences in conditioning history.

(*b*) Where plastics are rigidly attached to other materials, swelling pressures may act at the points of attachment and the materials will be correspondingly stressed. For example, in the experimental use of laminated plastics as aircraft wing covering, the sheet was attached by rivets to a metal framework. As a result of adsorption of moisture from the atmosphere, a force was exerted on the attachments, the sheet itself being stressed in compression. When this compressive stress exceeded the limits of elastic stability determined by the thickness and stiffness of the sheet, buckling (in one or more waves) occurred with serious aerodynamic consequences. Conversely, when moisture was desorbed, the sheet was stressed in tension which might in extreme conditions cause failure at the points of attachment.

(*c*) Where during transient conditions, the distribution of adsorbate in an article is not uniform, the portion having the higher concentration of adsorbate will be stressed in compression and balancing tensions will arise in the drier parts of the body. The corresponding bending moment will cause warping. Where relatively impermeable resins such as phenol-formaldehyde are used and where relative humidity varies cyclically, over a much shorter period than that required for equilibrium conditions to be reached, warping is probably the most likely cause of trouble.

Depending on the chemical composition of the polymer, the adsorbate may be a hydrocarbon or water, or it may even be a plasticiser added deliberately. Atmospheric water is the most common adsorbate, but it is well to remember that even if a polymer is unaffected by water, swelling by other solvents may have a considerable bearing on its use.

The authors' interest arises from an investigation into the mechanical properties of reinforced plastics, the swelling of which in presence of water constitutes the subject of this paper. Whilst, for convenience, values for moisture regain are generally quoted, it may be emphasised that it is the effects of the adsorbate on dimensions, strength, stiffness, resistivity and power factor, which are of technical importance.

The Swelling of the Component Materials.

Reinforced plastics consist of fibres bonded by resin, the latter enabling the high strength and stiffness of the fibres to be utilised. As a preliminary to consideration of the behaviour of the aggregate, the swelling phenomena of the component materials may be briefly discussed. With the exception of glass filaments and asbestos fibres, the most important reinforcements are of a cellulosic nature. Cotton or linen textiles and the paper-making fibres are, from a chemical standpoint, essentially cellulose, containing varying amounts of lignin, hemicelluloses and low molecular weight compounds. The moisture regain-vapour pressure relations of cotton, which is virtually pure cellulose, may be taken as representative, bearing in mind that the proportion of non-cellulosic material present in other fibres and also the extent of crystallinity influence the actual magnitudes of swelling. The sorption isotherms of cotton are sigmoid in form.[1] The value of moisture regain obtained at a given relative humidity is greater when equilibrium is approached from a higher relative humidity ("desorption") than from a lower relative humidity ("adsorption"), thus giving the effect of hysteresis. Dimensional movement in the longitudinal direction of the fibre, i.e. parallel to the direction of preferred orientation of the cellulose macromolecules, is negligible compared with that in a direction at right-angles, i.e., laterally. This behaviour has been associated with the observation that the X-ray spacings of crystalline cellulose are unaffected by sorption of water. It is, therefore, believed that water enters only the amorphous portions of the cellulose fibre, thereby displacing sideways the randomly arranged chains, but having little effect lengthwise.

It may be noted that replacement of cellulose in reinforced plastics by a non-hygroscopic fibre, such as glass or asbestos, although improving the resistance of the product to water, does not overcome this difficulty completely. In particular, the mechanical properties of such materials are adversely affected by adsorbed water.

---

[1] Urquhart and Williams, *J. Text. Inst.*, 1924, 15, T138.

The most important bonding agents in reinforced plastics are the condensation products of formaldehyde with phenol, urea and melamine. Reference may also be made to the diallyl polymers and silicone resins. Beyond the fact that all these polymers adsorb water (the silicones being the least hygroscopic) little is known about their equilibrium sorptions. The reason for this gap in knowledge is undoubtedly the fact that the permeability of these materials to water is exceedingly low, e.g. the moisture permeability constant[2] of phenol-formaldehyde varnishes is of the order of $0{\cdot}5 \times 10^{-8}$ g. cm./hour cm^2. mm. Hg. Hence measurements of the equilibrium values of sorption of such materials are unduly prolonged, unless very thin films are examined. An exploratory investigation of phenol-formaldehyde resins has been commenced and the desorption isotherm at 30° c. for a one-stage, ammonia catalysed, phenolic resin is given in Fig. 1. Thin discs (0·020 in. thick) were moulded from unfilled resin in pairs, one sample of each pair being dried over phosphorus pentoxide under vacuum. The other sample was saturated with water and then exposed to an atmosphere of known relative humidity.* After two years, the discs approached constant weight and from these data an approximate desorption isotherm was calculated. It will be observed that the relationship between moisture regain and relative humidity tends to be linear and less sigmoid in form than that of cellulose.

By use of a more rapid technique, it has been shown that the hygroscopicity of phenolic resins depends markedly on the nature of the phenol, on the catalyst used during condensation, and also on the degree of cure of the resin. In contrast to cellulose, moulded phenolic resin is mechanically isotropic and the linear swelling is the same in three directions at right-angles. It would appear that orientation during moulding is precluded by the extensive cross-linking of the phenolic macromolecule.

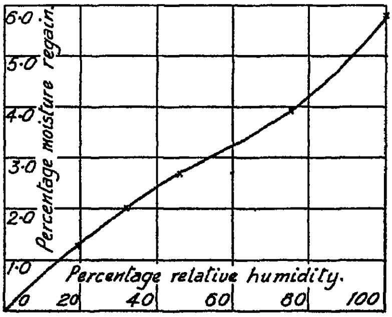


FIG. 1.—Desorption isotherm at 30° c. of unfilled phenol-formaldehyde resin.

The Swelling of Laminated Plastics.

In the manufacture of laminated sheet, the type of reinforced plastics with which the authors have been chiefly concerned, paper or fabric is impregnated with resin-solution, dried to remove solvent and moisture and cut to suitable size. The sheets of impregnated paper are arranged in a stack, like a pack of cards, and bonded under the influence of heat and pressure in a hydraulic press. A number of variables are involved in this process, the most important of which are discussed below.

Approximate adsorption and desorption isotherms for a sample of laminated sheet made from Manilla hemp paper and phenol-formaldehyde resin in aqueous solution are shown in Fig. 2 and 3. This laminate, which contained 50 % by weight of resin and was 0·020 in. thick, was bonded at 2000 lb. per sq. inch at a temperature of 145° c. for 30 minutes. Test pieces, 3 in. long × ½ in. wide, were cut from the laminate and dried over phosphorus pentoxide under vacuum for 6 weeks. The pieces were weighed and measured, and then exposed to atmospheres of known relative humidity for 6 weeks. After reweighing, the pieces were immersed in water, allowed to desorb at known relative humidities, and finally weighed.

[2] Thomas and Gent, *Proc. Physic. Soc.*, 1945, **57**, 324.

* Throughout this paper the values given for 100 % relative humidity were obtained by immersion in liquid water.

The data are expressed as % increases in weight and length on the dry values. Volume changes were not recorded, but it has been shown that, provided laminates are free from air-voids, the weight change is an adequate criterion of volume swelling. The sigmoid character of the moisture regain curve is again less pronounced than it is for cotton; there is but little evidence of hysteresis.

Changes in dimensions of laminated plastics consequent on sorption of water are greatly influenced by the geometry of the reinforcement and the manner in which the latter is arranged in the laminate. As a result of the more-or-less planar arrangement of fibres, these materials are not isotropic, possessing a greater tensile strength and Young's modulus in the plane of the sheet than through the thickness. When water is adsorbed, the resulting dimensional expansion is greatest in the least stiff direction. For example, the change in thickness may be some ten times greater than the expansion in any direction within the plane of the sheet. This latter movement is of the same order as that of plywood, and may differ in two directions at right-angles as a consequence of the directional properties of paper and fabrics. When, in preparation of a stack of impregnated sheets before bonding, alternate sheets are turned through a right-angle, the product is said to be "crossed" and the dimensional movements in the lengthwise and crosswise directions are equal.

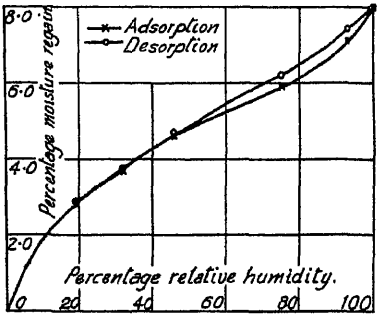


FIG. 2.—Moisture isotherms at 30° C. for sample of laminated sheet made from phenol-formaldehyde resin in aqueous solution.

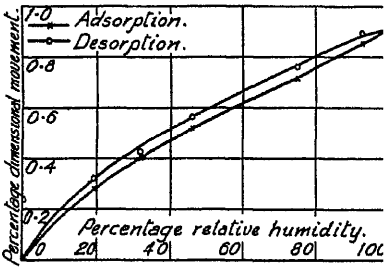


FIG. 3.—Dimensional movement at 30° C. in plane of laminated sheet made from phenol-formaldehyde resin in aqueous solution.

When, however, stacks are made up with all the sheets lying in one direction the laminated material is stronger and stiffer in the machine direction than in the crosswise direction and the dimensional movement in the latter direction is greater. The results given in Fig. 3 were obtained on a "crossed" laminate and are representative of any direction within the plane of the sheet.

It will be recalled that the movement of plywood in the plane of the sheet is very much less than the cross-grain movement of the component veneers. It appears that the cross-grain swelling of the face veneers is restrained by the core veneer, lying in a direction at right-angles to the face veneers. A similar effect has been observed with papers, where the dimensional movement of "crossed" laminates is usually less than the average of the movements in the machine and crosswire directions of uncrossed material.

When fabrics are used as reinforcement, the considerations described above are generally applicable, but the results are complicated by the tendency of twisted threads to shrink on adsorption of moisture. The dimensional movement decreases as the opportunity for the "guy-rope" effect to occur increases, i.e. with increase in the amount of air-voids; in fact, fabric reinforced plastics of low resin content and containing air-voids contract on wetting.

Factors in the Manufacture of Laminated Plastics Affecting Swelling.

The Type of Resin Solution.—As may be expected, the swelling of laminated plastics is profoundly affected by the type of resin solution used for impregnation of the reinforcement. The sorption isotherms are shown in Fig. 4 for a sample of laminated sheet prepared in a similar manner to that referred to in Fig. 2 and 3, except that in this case the phenol-formaldehyde resin was in alcoholic solution. Hysteresis is more pronounced for the alcoholic resin laminate and the latter possesses a lower moisture regain up to about 90 % R.H. Above this humidity, the aqueous laminate adsorbs less moisture and over a cycle typical of service conditions in Britain, e.g. 30 % R.H. to saturation, it is much less affected by moisture. The behaviour up to 90 % R.H. is consistent with the observation that unfilled resin *per se* from alcoholic solution possesses the lower moisture regain, but this does not account for the diminished swelling of the aqueous resin laminates above 90 % R.H.

The resin in aqueous solution impregnates the cellulosic reinforcement more thoroughly and probably penetrates the fibre-wall, where it may react chemically with the cellulose, thus reducing its hygroscopicity. On the other hand, the mechanism may be simply physical. The aqueous resin may be expected to enter microvoids on the surface of the fibre and thereby offer greater restraint to swelling. This view is supported by the observation that the aqueous boards possess the higher density.

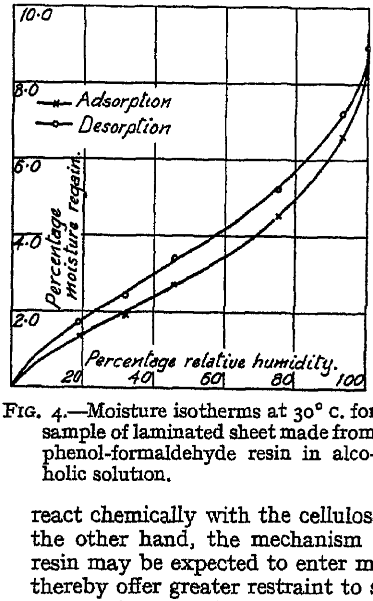


FIG. 4.—Moisture isotherms at 30° C. for sample of laminated sheet made from phenol-formaldehyde resin in alcoholic solution.

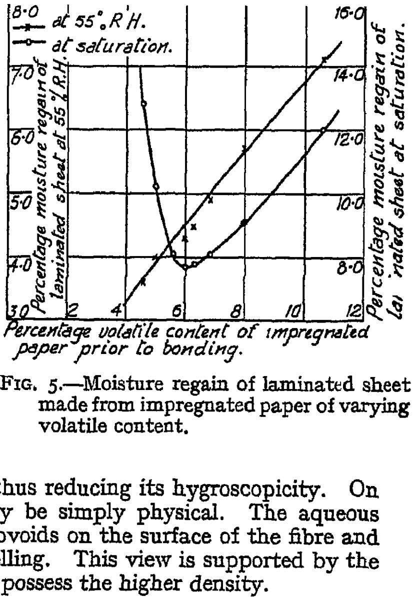


FIG. 5.—Moisture regain of laminated sheet made from impregnated paper of varying volatile content.

The Amount of Volatile Matter in the Impregnated Reinforcement Prior to Bonding.—After drying, the impregnated reinforcement still contains some volatile matter, namely residual solvent, hygroscopic water and potentially also, the water of condensation liberated during cure of the resin.

Fig. 5 illustrates the results obtained on laminates made from impregnated paper which was dried at 107° C. for varying periods and which, therefore, possessed varying volatile contents prior to bonding. The moisture regain of the laminates at 55 % R.H. will be seen to decrease progressively with reduction in volatile content. The amount of water adsorbed at saturation also decreases to a volatile content of 6 %, which, in this particular experiment, proved to be the minimum capable of giving a voidless product (see below).

This decrease which, of course, is also associated with a decrease in dimensional movement and is of importance in connection with the electrical and mechanical properties of reinforced plastics, has also been observed when the volatile content of impregnated paper or fabric is reduced simply by drying at room-temperature. It appears, therefore, that moisture retained in cellulosic laminates facilitates subsequent swelling. It is suggested that the resin offers mechanical restraint to the swelling of the fibre and that the drier is the fibre at the time when the resin hardens, the lower is the moisture content at equilibrium.

Air Voids.—It has been shown that to obtain maximum adhesion between resin and fibres, the air-voids in the paper or fabric structure must be completely filled with resin.[3]

Examination of the densities of the laminates referred to in Fig. 5 proved that air-voids were retained within material pressed from impregnated paper of volatile content less than 6 %. The air-voids are responsible for an increase in the amount of water adsorbed at saturation, but do not affect the moisture regain at 55 % R.H. Air-voids may also arise through the use of insufficient resin or insufficient pressure during bonding. Table I illustrates the latter effect on " crossed " laminates made from paper impregnated with 40 % of a cresylic resin.

TABLE I

LAMINATES BONDED AT VARYING PRESSURES

Bonding-Pressure (lb. per sq. inch).	Density (g. per cc.).	% Air Voids (calculated from density).	Percentage Changes on Desorption from Saturation to 32% R.H.			
			Weight	Volume	Length	Thickness
150	1·260	9·0	16·1	11·6	0·94	9·6
250	1·344	2·5	12·2	12·0	0·88	10·2
350	1·375	—	10·2	12·0	0·93	10·4
500	1·380	—	10·4	12·2	0·90	10·6
1000	1·382	—	10·2	12·3	0·87	11·3

The above results show that water, which enters voids, has little influence on the dimensional movement in the plane of the sheet. Further, the swelling of the resin-fibre aggregate may take place to some extent within the voids so that void-free laminates exhibit the greatest change in volume.

Resin Content.—The effect of resin content on swelling has been studied with a series of fabric base laminates made under carefully controlled conditions, resin content being the only variable. The measure of swelling given in Fig. 6(*a*) is the weight of water desorbed from saturation to equilibrium with 32 % R.H. expressed as a percentage of the weight at 32 % R.H. The dotted straight line indicates the amounts which would be expected if the laminates behaved as a simple mixture of two components. Fig. 6(*b*) shows the variation of Young's modulus with resin content for the same series. It will be noted that the amount of water desorbed is considerably less than that calculated, and that the greatest difference occurs at the resin content which gives the stiffest laminate. With one series of paper base laminates the relationship between swelling and resin content exhibited a minimum, i.e. the swelling of the laminate was less than that of either of the components.[4]

Fig. 7 shows the desorption isotherm for a laminate containing 50 % of resin (A), together with similar data for the two components (B and C). The dotted curve (D) gives the hypothetical behaviour of a mixture of the two components. These curves indicate that a reduction in swelling is obtained at all relative humidities.

---

[3] Cox and Pepper, *J. Soc. Chem. Ind.*, 1944, 63, 150.

[4] Pepper and Barwell, *ibid.*, 1944, 63, 321.

There is, therefore, *prima facie* evidence that combination of the two materials in a laminate introduces mechanical restraint which reduces swelling. Thus, swelling may be regarded as being resisted by the elastic stress in the material, equilibrium being reached when swelling pressure and elastic stress are balanced. If swelling pressure is regarded as being hydrostatic and the material is not isotropic, the resulting strains will vary in different directions, i.e. swelling will be least in the stiffest direction. Such a condition will, of course, give rise to shearing forces and, in the absence of the corresponding elastic data for the material, it is not possible to pursue the quantitative aspects of the problem.

The existence of hysteresis as shown in Fig. 4 places further difficulties in the way of quantitative treatment. Barkas has suggested that sorption hysteresis in gels is a consequence of mechanical hysteresis.[5] Comparison of the isotherms given in Fig. 2 and 4 shows that the laminated sheet made from aqueous resin solution, which is known to be more nearly Hookean in behaviour when stressed mechanically, exhibits less sorption hysteresis.

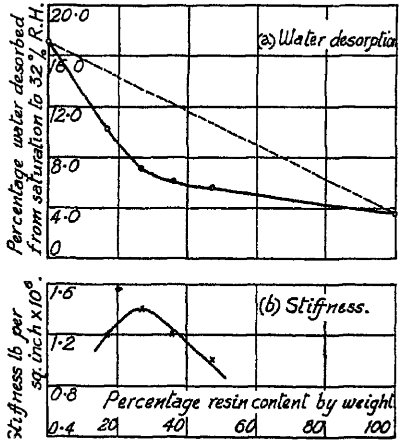


FIG. 6.—Water desorption and stiffness of fabric laminates of varying resin content.

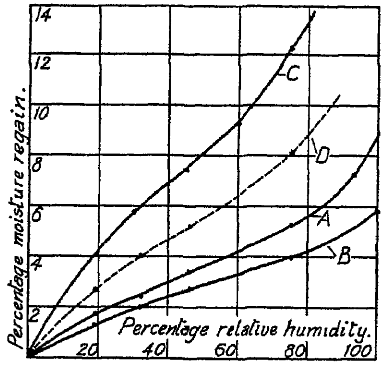


FIG. 7.—Desorption isotherms at 30° C. for a laminate and its components.

Conclusions.

The manufacture of reinforced plastics provides an example of the effect of combining two hygroscopic gels, namely cellulose and phenol-formaldehyde resin. The resin appears to reduce the hygroscopicity of the cellulose and in certain circumstances, the reduction may be mutual. While chemical effects cannot be excluded, e.g. interaction between resin or formaldehyde and cellulosic hydroxyl groups, the available evidence suggests that the effect is, in part at least, mechanical in character, the resin exercising a restraint on the cellulose. This view is supported by the fact that mere arrangement of the reinforcement, whereby the longitudinal stiffness of certain fibres is enabled to restrain the lateral swelling of other fibres, as, for example, in " crossed " boards, may bring about a significant reduction in the swelling of the composite.

Further, the observed reduction in swelling of laminates resulting from decrease in the volatile content of the impregnated reinforcement before bonding, can be most satisfactorily explained on the basis of increased mechanical restraint.

[5] Barkas, *Swelling Stresses in Gels*, 1944, 53.

Summary.

The behaviour of reinforced plastics in presence of water has been considered from the standpoint that both components, resin and cellulose, are swelling gels. Sorption isotherms of a sample of unfilled phenol-formaldehyde resin and of two types of laminated plastics are given. The influence of certain manufacturing variables on swelling is outlined.

It is concluded that, in laminated plastics, the swelling of the components is reduced, and there is *prima facie* evidence that this reduction is in part due to mechanical restraint operating mutually on resin and fibre.

The work described was carried out in the Engineering Division, National Physical Laboratory, and in the Chemical Research Laboratory, on behalf of the Director-General of Scientific Research (Air), Ministry of Supply, by whose permission this paper is published.

The authors desire to acknowledge the assistance rendered by Mr. D. J. McCauley of the Chemical Research Laboratory, who performed the experimental work.

Department of Scientific and Industrial Research,
Engineering Division,
National Physical Laboratory,
Teddington, Middlesex.

Chemical Research Laboratory,
Teddington,
Middlesex.

*GENERAL DISCUSSION**

Dr. W. W. Barkas (*Princes Risborough*) said: To see the progress made in the understanding of the swelling of reinforced plastics as a result of the work of the present authors and others, it is interesting to look back only a few years to the early days of the war. It was then commonly assumed:

(1) that the reduction in sorption was due to a real reduction in the hygroscopicity of the swelling material (paper, wood, etc.) by some undefined chemical reaction with the plastic;

(2) that the resin (plastic) component was non-hygroscopic and non-swelling, so that the sorption could be put down entirely at the door of the cellulosic component;

(3) that the swelling was equal to the volume of liquid adsorbed, so that either the swelling or the weight increase could be used as a measure of the other.

The authors have studied all these points and have reached important conclusions regarding them. With respect to (3) they find that the relationship holds for their material due to the absence of voids, but it is worth pointing out how necessary it is to determine this. For example, in natural wood:

(*a*) the presence of voids would lead one to expect that in free swelling the external volume increase in a wooden block would be about three times the volume of water adsorbed by the cell wall, but

(*b*) it is found that due to anatomical restraints imposed on the cell wall much of the swelling occurs inwards into the voids, so that in reality the external swelling is only about 1·5 times the volume of water adsorbed.

Mr. F. I. G. Rawlins (*London*) said: The paper by Bardwell and Pepper contains the core of the problem confronting curators of paintings, especially " old masters " executed upon wood panels. Briefly, the structure of these, often priceless, objects, is a four-fold stratified one arranged as follows (starting from the back):

---

* On preceding paper.

(*a*) the support—wood, canvas or paper (occasionally metal);
(*b*) the ground—usually a calcium water paint, sometimes Pb salt;
(*c*) the paint-film—an aggregate, pigment plus film substance;
(*d*) the surface-film—varnish, resin, wax, etc.

Whereas (*a*) differs fundamentally from (*b*), (*c*), (*d*), these latter three can be, as a first approximation, grouped together as "superincumbent layers" (*sl*). We thus have a two-fold system [(*a*), (*sl*)] to consider. Not only the materials, but the thickness of them, too, may be widely different—e.g. (*a*) may be as much as 50 mm. thick (for a heavy Italian panel of poplar wood), and (*sl*) about 1 mm. The former, of course, is anisotropic in bulk, the latter, anisotropic over comparatively small areas (mosaic-structure) due, naturally, to the design exhibited by the paint-films. This local anisotropy seems to have little effect upon the stability of the whole complex, cancelling out statistically unless the pigments used are exceptionally incompatible. This "horizontal" variation in property is, however, probably not the most important. When films such as occur in the studio are first applied they are composite in that the "top" one is rigid, and the lower one (that in contact with the surface being painted) is more or less plastic. It takes a long time, in the usual sense of the word, say a year, for the rigid layer to grow in depth sufficently to convert the whole covering to such a condition, but it is quite certain that in the great majority of "old masters" (*sl*) is practically rigid. (Exceptions will be mentioned later.) This is probably the cause of a somewhat remarkable phenomenon observed by the Committee set up may years ago to study the stresses and strains—i.e. shrinking and swelling—suffered by pictures as a result of variations in relative humidity. The effect in question is a well-known one, namely that the changes in dimensions taking place at the front of a panel are remarkably small, much too small in fact to be accounted for by (*sl*) acting as a protective coating against atmospheric fluctuations. (In any case, the permeability of most film materials used in the arts is quite high.) Thus, it appears that what happens is that (*sl*) clamps down upon (*a*) with very considerable force, inhibiting shrinking and swelling of the latter, until some critical value is reached (perhaps as a result of relative humidity cycles without any gross mechanical ill-treatment at all), when (*sl*) gives up the unequal battle, and sudden collapse occurs. This is indeed a state of affairs by no means unknown, the abrupt failure of a painting, first in the ground (*b*) and then in the paint-film (*c*). Experiments are about to be started to test this with much greater refinement than was feasible some fifteen years ago. Some work done in U.S.A. during the War is also significant. Amongst other factors to be considered is that of relative humidity at "near surface" points.

Two complexities, indicated already, should be mentioned in conclusion. One is the assumption that the system (*sl*) can be assumed to be rigid. Only too obviously, in some instances this is not so. Cases are on record where an undercoating of oil has been found, and on which —so to say—the whole superincumbent system tends to slide. Another is the habit of Sir Joshua Reynolds in his addiction to bitumen. This substance never really solidifies (so far as studio technique goes); in fact it flows slowly, like a glacier, for ages, making pictures of this master the despair of anybody whose responsibility it is to maintain them in good condition. Work has been done on the rheological characteristics of some artists' materials, but much remains to be done.

Finally, the mechanics of the surface-film (*d*) may be more severe than is commonly imagined; or at least its interaction with its neighbour (*c*) may be. Resinous varnishes should be kept thin (if used at all). Otherwise they may exert such a "pull" upon the paint-film below as to detach portions of it in curly scales. Naturally, repeated application of varnish to paintings (a common practice in the past) is most dangerous from this aspect.

Altogether, in recent years, a new science called " Picture Mechanics " has come into being to study all these effects. It requires the very careful and discriminating application of " properties of matter " and thermodynamics to the complex systems discussed in this note.

Dr. K. W. Pepper (*Teddington*) said : The effect described by Mr. Rawlins provides another example of inhibition in swelling and shrinking brought about by mechanical restraint. Sudden collapse of the ground or paint-film by fatigue is the likely result when reversal of strain is repeated many times. We have, indeed, predicted such a failure in reinforced plastics, should these materials be exposed to sufficient cyclic changes in humidity, although of course the time required to produce a sufficient number of reversals in condition has precluded confirmation by experiment.

THE ADSORPTION-SWELLING OF SEVERAL CARBONACEOUS SOLIDS.

BY F. A. P. MAGGS.

Received 19th August, 1946.

It has been established by Bangham and his co-workers[1] that the percentage linear swelling (x) of a charcoal block during adsorption is directly proportional to the free energy decrement ($-\Delta G$) of the block as estimated from the pressure-adsorption relations by use of the Gibbs' equation.[2]

$-\Delta G$ was identified as a *surface* free energy change on the grounds that the same stoichiometric relations were found to hold between ΔG and the quantity adsorbed, as between the surface tension lowering and surface concentration of the same adsorbates on mercury surfaces.[3] It follows that the charcoal used was characterised by a definite surface area (Σ) per g. and that the free energy may be regarded as a surface pressure (F dyne/cm.).

$$x = -K\Delta G = KF\Sigma. \quad \ldots \quad (1)$$

The constant K, Bangham pointed out, was probably related to the elastic constants of the charcoal.

The physical significance of K can be more clearly seen if we make simplifying assumptions as to the structure of charcoal. We may, for example, suppose that 1 g. of charcoal (density ρ, g./cc.) takes the form of a long thin cylinder of radius r and curved surface Σ; if, then, a surface force of $2\pi rF$ acting on a cross-sectional area πr^2 gives rise to a linear extension x %, the ratio stress/strain is given by

$$E_A = 100\ \rho F\Sigma/x.$$

The same relation holds for a laminar model. However, the result is not in general independent of the geometry of the model (although other models give similar relationships). Moreover, the modulus E_A cannot be strictly identified with Young's modulus E_M since the volume contraction associated with a contraction normal to the applied tension is not necessarily the same in the two cases. On the other hand, E_A and E_M would not, on theoretical grounds, be expected to differ by as much as an order of magnitude, and in favourable cases, they may well approach equality.

---

[1] Bangham and Fakhoury, *J. Chem. Soc.*, 1931, 1324; Bangham, Fakhoury and Mohamed, *Proc. Roy Soc. A.*, 1932, 138, 162.

[2] Bangham, *J. Chem. Soc.*, 1931, 1324; *Trans. Faraday Soc.*, 1937, 33, 804.

[3] Bangham, Fakhoury and Mohamed, *Proc. Roy Soc. A.*, 1934, 147, 152.

The main purpose of the experiments to be described was to measure E_A and E_M for several carbonaceous solids—charcoal blocks, monoliths of coal, and compressed coal powder artefacts after heating—in order to gain information as to their inner structure by investigating the applicability of equation (1).

Experiments with Charcoal.—For the measurement of E_A, a rod of unactivated willow charcoal taken from the stock used by Bangham and Razouk[4] was supported in a silica extensometer and thoroughly degassed whilst being heated. Methanol was then distilled into the containing vessel and the expansion (x_L) on complete immersion at 25° c. was measured. The corresponding free energy of immersion was not measured directly. Instead, the heat of immersion ($-\Delta H_L \Sigma$ cal./g., H being the total energy per cm.^{2}) was determined and Bangham and Razouk's value for the ratio of $\Delta H_L \Sigma / \Delta G_L$ was used.* These experiments gave: $-\Delta H_L = 14{\cdot}5$ cal./g.; $-\Delta G_L = 7{\cdot}25$ cal./g.; $x_L = 0{\cdot}713$ cal./g.; $E_A = 71 \times 10^9$ dyne/cm.^{2}. This value of E_A compares reasonably well with that of 110×10^{10} dyne/cm.^{2} calculated from the corresponding data of Bangham and Razouk for their charcoal IV.

E_M was determined by measuring with a dial gauge the compression of the rod between two parallel plates. Stress-strain curves are shown in Fig. 1. An apparent modulus of 22×10^9 dyne/cm.^{2} was calculated from these curves. Further experiments showed that the modulus determined in extension was the same as that in compression. Whilst the changes in length are considered accurate to 5×10^{-3} % of the length, some uncertainty attaches to the true value of E_M, for the following reasons. Since the fragile nature of the charcoal precluded machining, the cross-sectional area is open to an error of 10 %. Moreover, the greater part of the pore volume consists of visible sap ducts running parallel to the direction of the applied stress, and whereas Meehan[5] found charcoal to swell isotropically during adsorption, it was not practicable to measure E_M in other directions. The apparent modulus was, therefore, multiplied by the ratio of the density determined by helium displacement to the block density, on the assumption that the pore volume consisted mainly of parallel sap ducts; other simple assumptions, such as random distribution of the pore volume, give lower values for the "true" modulus.

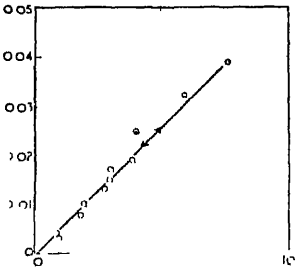


FIG. 1.—Stress-strain curve for charcoal; ordinate: strain, percentage compression; abscissa: stress, kg./cm.^{2}

The experimental results were: apparent Young's modulus = 22×10^9 dyne/cm.^{2}; block density = 0·34 g./cc.; density in helium = 1·63 g./cc. †; whence $E_M = 100 \times 10^9$ dyne/cm.^{2}.

The agreement between E_M and E_A is thus clearly within the margin of uncertainty arising from the calculation of the former.

To examine whether the expansion or immersion had extended the specimen beyond the region of Hooke's law, a further compression stress-strain curve was determined with the charcoal immersed in methanol. The stress-strain relations were unaltered by wetting.

---

* Calculated by them from the experimental data by use of the Gibbs-Helmholtz equation:

$$-\Delta H_L = \frac{1}{K}(x_L - T dx_L/dT).$$

† Measured by R. E. Franklin.

[4] Bangham and Razouk, *Proc. Roy. Soc. A.*, 1938, **166**, 572.

[5] Meehan, *ibid.*, 1927, **115**, 199.

No account has been taken so far of the work expended on swelling the charcoal by the surface pressure. For a solid undergoing deformation in accordance with Hooke's law, the work of swelling in one direction is equal to $Ex^2/(2 \times 10^4)$. For this charcoal, the work of swelling amounts to less than 1 % of the wetting free energy change, and is insignificant.

Experiments with Coals and Heated Coal Artefacts.—It was observed by Lea [6] that coal swells in water, and since this is most probably an adsorption and not a solution effect, a comparison of coal swelling with charcoal swelling is of considerable interest from the structural aspect.

The coals used in the following experiments include a lignite, but no anthracites.

Coal Monoliths.—Three series of experiments have been carried out, using methanol as the adsorbate.

(i) An adsorption-expansion isotherm on a coal monolith was measured. [*]
(ii) The immersion swelling and heat of wetting of a number of evacuated coal monoliths were determined. [*]
(iii) The saturation adsorption of methanol on a number of coals was determined.

The results are summarised in the following statements.

1. No clear-cut relation between the swelling and $-\Delta G$ (determined by integration of the isotherm), such as was found for charcoal, was observed; hysteresis phenomena complicate the issue. On the other hand the ratio $x/(-\Delta G)$ shows only a two-fold variation and is more constant in desorption than adsorption.

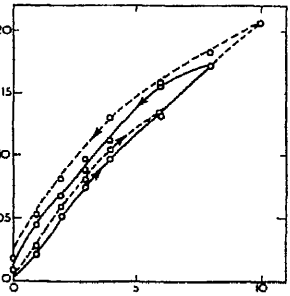


FIG. 2.—Stress-strain curve for coal monolith; ordinate: strain, percentage compression; abscisa: stress, kg./cm.^{2}

2. The mechanical stress-strain curves for several coals are not linear and show hysteresis (see Fig. 2).

3. Swelling anisotropy is observed with coals, the swelling in the direction perpendicular to the "bedding" plane being greater than that in the bedding plane; the ratio shows surprisingly little variation about a value of 1·7. The anisotropy of swelling is in direct contrast with X-ray, optical and thermal expansion data, where little or no anisotropy is observed.

4. E_M is greater than E_A the ratio varying from 2 to 11; some values of E_M and E_A are given in Table I. A rough proportionality exists between heat of wetting and immersion swelling, so that E_A does not vary greatly from coal to coal.

TABLE I.

Coal.	A.		B.		C.		D.	
Orientation in bedding plane	‖	⊥	‖	⊥	‖	⊥	‖	⊥
$E_M 10^9$ dyne/cm.^{2} . .	20	40	13	20	30	45		20
$E_A 10^9$ dyne/cm.^{2} . .	5·7	4·2	5·1	3·6	6·9	3·8	9·5	4·8

5. Whilst E_A is greater when measured parallel than it is normal to the bedding plane, the converse holds for the few values of E_M, as measured in compression. The discrepancy between E_A and E_M is thus much less in the direction of the bedding plane than normal to it.

---

[6] Lea, *Fuel in Science and Practice*, 1928, 7, 430.

[*] For description of apparatus and method see "Conference on Ultrafine Structure of Coals and Cokes", B.C.U.R.A. 1943.

6. The expansion of coal monoliths is directly proportional to the volume of vapour adsorbed, both over the course of an isotherm for a single coal, and for the saturation adsorption and immersion swelling for a number of coals. The *volume* swelling is about half the volume of methanol adsorbed, calculated as normal liquid. This relationship contrasts strongly with that for charcoal, where the volume swelling is but a small proportion (less than one-tenth) of the volume of adsorbed liquid.

It must be observed that whilst the porosity of charcoal is *ca.* 80 % of the block volume, the pore volume of those coals which show a large immersion swelling and which adsorb strongly is only *ca.* 16 % and for coals which swell but little and adsorb but little vapour at saturation, only *ca.* 3 %.

A comparison with R. E. Franklin's density data for the same coals shows the following quantitative relationship between swelling and adsorption to exist:

(volume of MeOH adsorbed at saturation) — (total volume contraction of system, i.e. surface compression) = (volume swelling of the block on immersion) + (pore volume of the dry coal).

All four quantities were measured in independent experiments.

7. The work of swelling—although difficult to estimate on account of the departures from Hooke's law, and the uncertainty in the value of the elastic constants—varies from coal to coal. As a fraction of $-\Delta G_L$, the work of swelling for coals where x_L has a low value, is smaller than in the case of charcoal, whilst when x_L is large, higher values are obtained (probably as much as ten times greater than those for charcoal).

Experiments with Artefacts.—The powder compacts were prepared from a weakly caking coal by cold-pressing coal powder (— 72 B.S.S.) at 6 tons/in.^{2}; a number of these specimens were then heated in a stream of nitrogen to a series of temperatures.

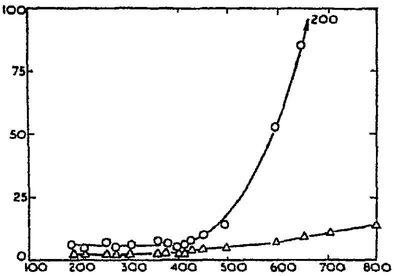


FIG. 3.—Variation of E_A and E_M of artefacts with temperature of preparation. Upper curve refers to E_A, lower curve to E_M; ordinate: modulus, 10^9 dynes/cm.^{2}; abscissa · temperature, °c.

E_A was estimated for each artefact from the immersion swelling and the heat of immersion in methanol; Young's modulus was also determined in compression. E_A and E_M are compared in Fig. 3 by plotting both moduli against the temperature of preparation of the artefact.

The results are summarised as follows

1. E_A is independent of temperatures of preparation lower than 450° c.; its value lies between the "parallel" and "perpendicular" values of the unheated monolith of the same coal.

2. The stress-strain curves are not linear; (the secant modulus has been used in Fig. 3). The values of E_M increase smoothly with increasing temperatures of preparation. At low temperatures, the value is much smaller than that for the monolith of the same coal.

3. Values of E_M are consistently smaller than those of E_A, and above 400° c., in the range where marked structural changes occur, E_A/E_M rises sharply.

4. E_M determined for artefacts prepared between 300° and 350° c. (where the coal structure is changed but little) was found to be *decreased* to about half or one-third of its original value by determining the stress-strain relations (in compression) under methanol (see Fig. 4). A value close to the original was *regained* by pumping off the methanol.

The conclusion to be drawn from these experiments is that whilst

E_A measures a property of the fine structure of the coal, E_M is influenced by the particulate nature of the artefacts.

Mechanism of Coal Swelling.—Van der Waals forces acting over areas of contact are, to a large extent, responsible for the cohesion between the structural units of coal. The term "area of contact" however, requires qualification.

Orowan[7] has shown that Griffith cracks, which may initially form part of the area of contact, can in the presence of an adsorbed film become large enough to weaken the structure; the extent to which this occurs is a function of the surface pressure F.

A modification of Orowan's theory appears to be applicable to coal. Three types of surface may be distinguished in coal: (1) surfaces held together by van der Waals forces at the molecular equilibrium distance; (2) surfaces enclosing the pore volume measured by helium displacement; (3) surfaces within the range of each others van der Waals forces, but held at a distance of separation greater than the equilibrium position by the structural rigidity of the solid. Discontinuities of the last type would constitute Griffith cracks.

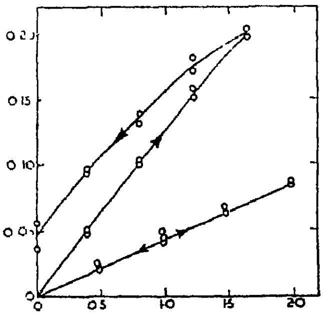


FIG. 4.—Stress-strain curve for artefact prepared at 300 °C.; lower curve: dry artefact; upper curve: artefact immersed in MeOH; ordinate: strain, percentage compression; abscissa: stress, kg./cm.^{2}

In the presence of an adsorbed film the coal swells in the manner described for charcoal. Superimposed on this swelling, however, is that due to an opening of Griffith cracks, which will be greater the greater the surface pressure.

Two effects might be expected to result from this. Firstly, a weakening of the structure would result on adsorption; secondly a simple relation between volume swelling and volume adsorbed should be observed when small cracks are filled; the limited pore volume of the lump coal would also lead to a similar result.

The surfaces of type (1) are not separated by adsorbed films for the swollen coals still show considerable resistance to deformation; moreover, peptisation does not occur.

Summary.

It is concluded that with charcoal (as opposed to coal) the swelling relations are determined by its obedience to Hooke's law, the dimensional change being proportional to the pressure of the films formed at its inner surfaces. With coal, the stress-strain relations are complex; and though there is much to support the view that the film pressure here also provides the swelling force, clear-cut evidence of a simple relation between the adsorptive and mechanical properties is, on the whole, lacking. A simple volumetric relation is found between swelling, adsorption and pore volume.

The British Coal Utilisation Research Association,
13 Grosvenor Gardens, London, S.W.1.

---

[7] Orowan, *Nature*, 1944, 154, 341.

THERMAL EXPANSION OF COALS AND CARBONISED COALS.

BY D. H. BANGHAM AND ROSALIND E. FRANKLIN.

Received 19*th August*, 1946.

The coefficient of thermal expansion of a solid, being determined by the vibrational amplitudes of an orchestra of atomic—and larger—oscillators, should depend on its molecular constitution rather than upon the physical structure on a colloidal scale of magnitude. With the important reservation discussed in the next section, it can be regarded as an intra-micellar property rather than one concerned with micellar surfaces and inter-micellar contacts. Its variation as between different specimens of coal was expected to throw light on a number of obscure questions, as, for example :

1. whether or not the anisotropy of coal—the mechanical properties of which vary markedly with orientation with respect to the " bedding " plane—extends to the atomic (or molecular) dispositions within the micelles;
2. whether or not there is any marked trend in the size or nature of the dynamic unit of structure with increasing " maturity " of the coals;
3. whether the important changes in the technical qualities of coal engendered by preheating below 300° c. (such as the suppression of " caking properties ") are associated with intramicellar chemical change, or confined to the micellar surfaces and contacts. Aside from the other forms of chemical change, any variation in the size or number of the dynamic units—whether brought about by condensation, polymerisation or cross-linking—would be likely to reveal itself in a dependence of the expansion coefficient upon temperature of pre-treatment.

Thermal Coefficient of Adsorption Swelling.—The presence of films at the micellar surfaces introduces a complication which is liable to obscure the issues. Confirming an earlier observation, Bangham, Fakhoury and Mohamed[1] found that for a given quantity (s) of a lower alcohol adsorbed on charcoal, the percentage linear expansion (x_s) was directly proportional to the absolute temperature T. The effect upon the thermal expansion coefficient (α) of the charcoal is considerable.

Degassed charcoal . . . . $\alpha = 4{\cdot}5 \times 10^{-6}/°$ c.
Charcoal with nearly saturated film of lower alcohol
(constant adsorption) . . . . . . $\alpha = 33 \times 10^{-6}/°$ c.

The relationship $x_s \propto T$—interesting theoretically since it emphasises the essentially dynamic nature of adsorption swelling—is to be expected only when the equation

$$x = \lambda F$$

(where F is the surface energy lowering) is obeyed and the constant λ is independent of temperature in the range considered. Direct proportionality of x to T is to be expected, also, only when the differential heat of sorption $(\partial H/\partial s)_T$ is independent of the surface concentration (as for methanol adsorbed on charcoal).

---

[1] *Proc. Roy. Soc. A.*, 1934, **147**, 160.

The relations

$$T(\partial F/\partial T)_s = F \text{ where } (\partial H/\partial s)_T \text{ is constant,}$$
$$T(\partial F/\partial T)_s > F \text{ where } (\partial H/\partial s)_T \text{ increases with } s,$$
$$T(\partial F/\partial T)_s < F \text{ where } (\partial H/\partial s)_T \text{ decreases with } s,$$

are readily derived thermodynamically.[2]

Unlike $(\partial F/\partial T)_s$ and $(\partial x/\partial T)_s$, to which the above generalisations apply, the coefficients $(\partial F/\partial T)_p$ and $(\partial x/\partial T)_p$ (where p is the pressure of vapour in equilibrium with the adsorbed film) are *negative*, on account of the vaporisation of the films.

In practice the conditions for measuring $(\partial x/\partial T)_s$ are difficult of achievement. The dimensional changes undergone by a rod of charcoal (or coal) containing moisture, when heated at a steady rate, are determined by the relative rates of inward heat transfer and of outward mass transfer of moisture. What usually happens is that the thermal expansion is at first markedly greater than for the dried material; thereafter it falls off and there is shrinkage over a relatively small temperature range followed by renewed expansion when the films have evaporated. On cooling again in a dry atmosphere contraction is continuous.

Method of Experiment.—To avoid the obscuring effects arising from the adsorption expansion the following method of experiment was adopted The specimen was heated in dry nitrogen and maintained at (say) 100° c. until the shrinkage rate became insignificant. The temperature was then lowered by stages with readings every few degrees and again raised slowly to 100° c with further readings It was then raised rapidly to (say) 150° c, and the whole procedure repeated. The maximum preheat temperature was thus raised by *ca* 50° c. in successive experiments.

MATERIALS.—Three " bright " coals of different carbon content were used, these being in the form of rod-shaped monoliths or cold-pressed artefacts (without binder) as ease of preparation dictated. For making the artefacts the coal, ground to pass a 72-mesh B.S. sieve, was pressed at 12,000 lb./in.^{2} for 1 minute.

The coals were :

Coal K, a " bituminous " coal of high oxygen content, from which both monolith and artefact rods could be produced.

Coal F, a " bituminous " coal of high carbon content used technically for coke manufacture; this coal being very friable, only artefacts could be used.

Anthracite C, from which could be obtained monoliths (both parallel and normal to the " bedding plane ") but not artefacts without binder.

The analyses of these coals are set out in Table I.

TABLE I.

	Ash %.	Per cent. (Parr's basis).				Heat of wetting in MeOH (cal./g).
		Carbon.	Hydrogen.	Nitrogen.	Oxygen + errors.	
K	1·0	82·4	5·1	1·9	10·6	16·6
F	4·1	89·7	5·0	1·8	3·5	2·0
C	2·1	94·2	2·9	1·2	1·7	7·6

Apparatus.—The samples, 64 mm. in length, were mounted between the jaws AA′ of a silica optical-lever extensometer illustrated in Fig. 1. The silica plates DD′, lightly ground and accurately parallel, are at such distance apart that when the brass roller R (5 mm. diameter) is placed between them the springs CCC′ are slightly strained. The surface of R is lightly ground and at one end there is a mirror the reflecting surface of which is coplanar with the axis of R. The extensometer is supported in a Pyrex jacket, one end of which is closed by a sealed-in optically plane window and protrudes from the furnace. Light from a lamp is projected on to the mirror and focussed on a scale at 125 cm. distance. The scale could be read to 0·1 mm., and the overall magnification obtained was × 500.

---

[2] *Trans. Faraday Soc.*, 1937, 33, 805.

Inside the furnace, and insulated from the nichrome windings, was a heavy brass block bored to fit closely the jacket of the extensometer. This gave within the jacket a zone 13 cm. long within which the temperature did not differ by more than 0·5° c. from that of the block.

Temperature measurements were made with Pt/Pt-Rh thermocouples, one located in a hole in the furnace block and another at the centre of the specimen. Thermocouples were incorporated in the artefacts at the time of pressing.

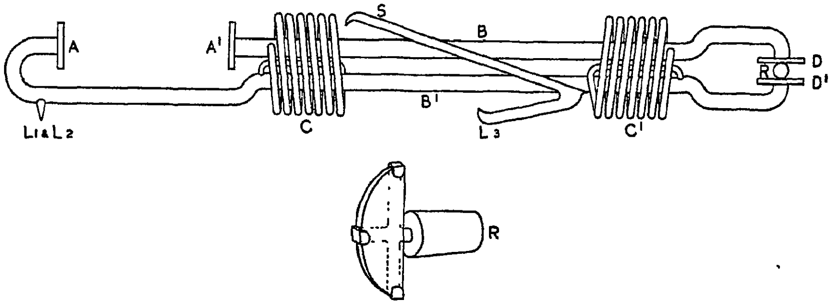


FIG. 1.

Throughout each experiment a steady stream of dry oxygen-free nitrogen was passed through the jacket at 18 litres/hr.

Results and Discussion.

Abnormal Expansions Associated with First Heating.—During the initial relatively rapid heating of Coal K to each new maximum temperature the value of α exceeded that observed during the subsequent cooling and reheating over the same temperature range. The data in Table II illustrate this point; the abnormal coefficients are shown in brackets. There can be little doubt that these abnormalities arose from the presence of adsorbed films and are the counterpart of the effects noted with charcoal. As the temperature of the specimen is raised steadily, the period of abnormal expansion is followed rather sharply by one of *shrinkage*. In the lowest temperature range the film which volatilises is readily identifiable as moisture, but this does not necessarily apply to the higher temperatures.

TABLE II.—COAL K.

Range (°c.).	$\alpha \times 10^6/°$ c.
20- 55	= 32(41)
55-100	= 39(45)
100-150	= 43(49)
150-200	= 48(53)

The low thermal conductivity of coal precluded rapid uniform heating and further examination of this phenomenon was not attempted. The abnormal thermal expansion was less conspicuous in the case of Coals F and C which have much smaller adsorbing surfaces than K.

Irreversible Shrinkage.—The isothermal shrinkage of Coals F and K in the range below 400° c. is plotted as a function of temperature in Fig. 2. The results have no absolute significance but depend on the procedure outlined. In the range below 100° c. the shrinkage of the Coal K is much more marked than with the much less hygroscopic Coal F.

We are indebted to J. Sandor for an observation which confirms that the shrinkage of coal K below *ca.* 100° is readily reversed by allowing the sample to take up moisture. This is no longer true of the shrinkage at higher temperatures.

Reproducible Thermal Expansion.—The following generalisations summarise the results obtained with specimens after the effects of abnormal expansion and shrinkage had been eliminated by preheating.

1. The expansion coefficient for each temperature range was practically independent of the preheat temperature below 300° c.

2. Within each temperature range, α for the two bituminous coals was independent of: (*a*) "rank"; (*b*) whether artefact or monolith; (*c*) pressure applied during fabrication (of artefacts); (*d*) orientation with respect to the bedding plane (of monoliths).
3. For the bituminous coals F and K, α increases steadily and markedly with temperature (Fig. 3): $\alpha_{30°}=33\times10^{-6}/°$ c.; $\alpha_{330°}=60\times10^{-6}/°$c.

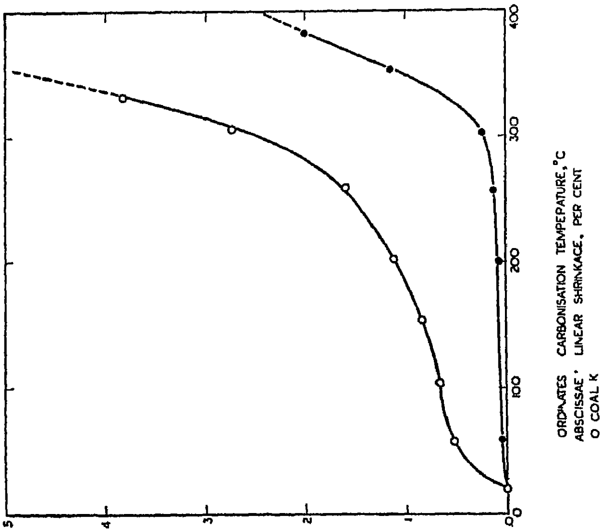


FIG. 2.

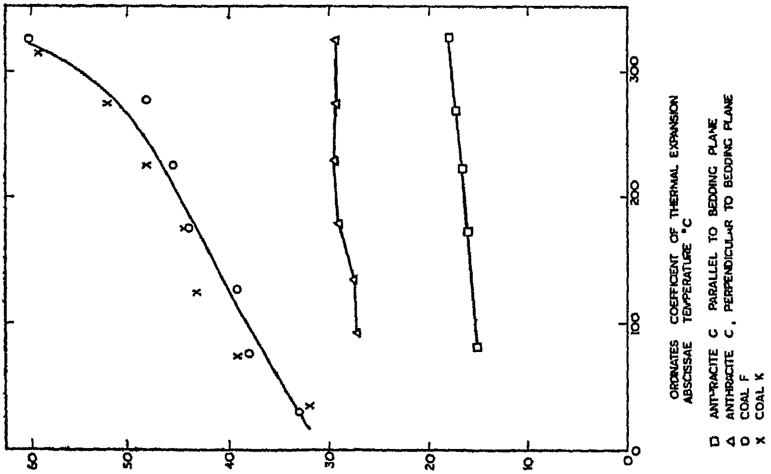


FIG. 3.

4. When the two bituminous coals were preheated above 350° c. (for F) and 325° c. (for K), α at all lower temperatures was greatly diminished by this treatment (Fig. 4 and 5), though it remained temperature-dependent.
5. For carbonisation temperatures above 600° c. for coal F and 550° c. for coal K the volume expansion coefficient 3α falls well below the value recorded by Riley and Nelson[3] for macrocrystalline graphite.

---

[3] *Proc. Physic Soc.*, 1945, 57, 477.

(In Fig. 4 and 5 the broken line represents one-third the volume coefficient calculated from Riley and Nelson's data.)

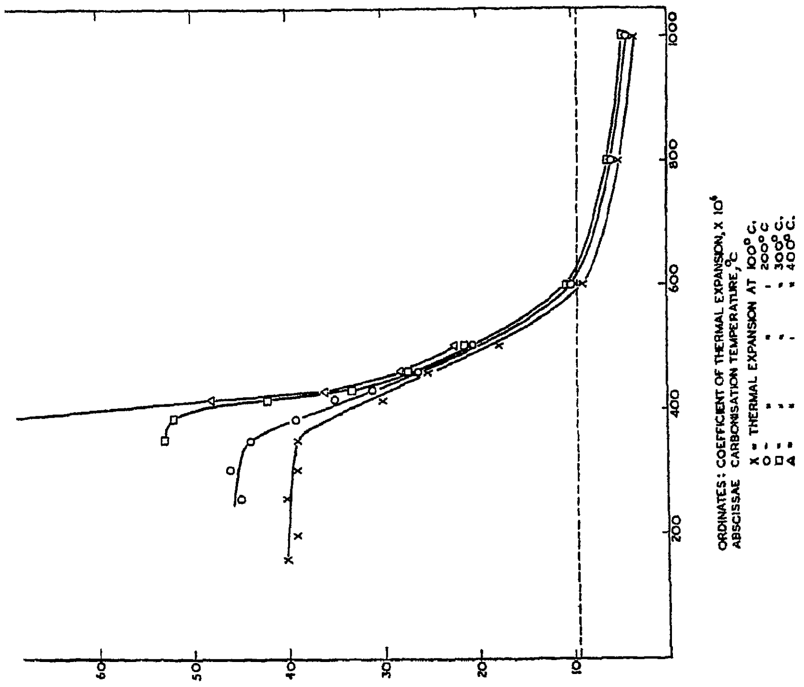


FIG. 4.

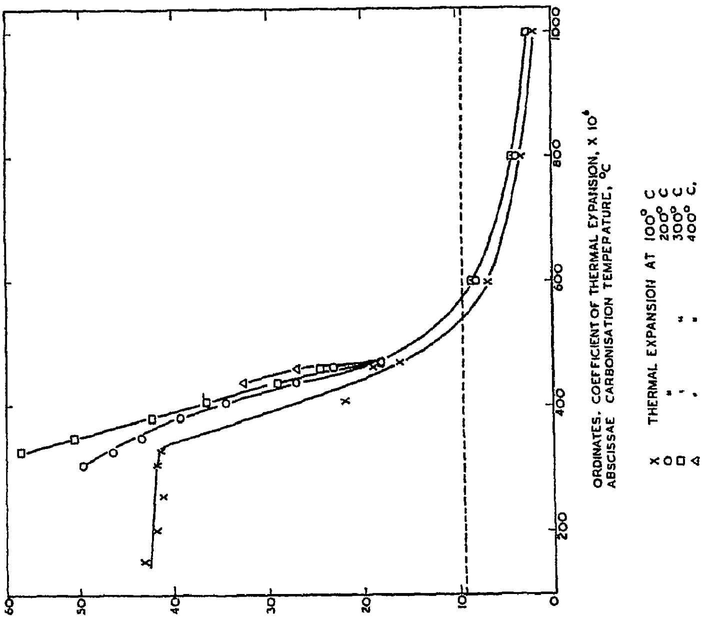


FIG. 5.

6. The high carbon bituminous coal F yields a coke for which α is significantly greater than that from the high oxygen coal K.

7. For the *anthracite* C, α varies much less with temperature but is markedly dependent on orientation (Fig. 3). Perpendicular to the bedding plane its value is remarkably close to that reported by Riley and Nelson for graphite normal to the layer planes; but whereas graphite between 0° and 400° contracts in directions parallel to layer planes, α for anthracite parallel to the bedding plane is positive and increases with temperature.

Discussion.

The values of α for the coals F and K far exceed those characteristic of materials having a continuous covalent network structure, and this suggests that forces of the van der Waals type play an important part in determining their state of aggregation. These relatively high values are retained when the coals are heated to 385° C. (F) and 350° C. (K); below these temperatures there is no evidence of polymerisation, cross-linking, or deep-seated chemical change such as may account for the diminution which sets in sharply at higher preheat temperatures. The two coals behave remarkably alike but this does not apply to their carbonised residues.

At room temperature α for coals F and K has a value $33 \times 10^{-6}/°\text{C}$. which is low in comparison with most organic high polymers, but comparable with that for some vulcanised rubbers. Whereas many organic polymers (and some glasses) show evidence of a second-order transition, characterised by a sharp increase in the coefficient from one nearly constant value to another, the coals show a *gradual* increase in α (indicative of an increase of internal mobility) over a wide temperature range.

It is indeed remarkable that the values of α and of $d\alpha/dT$ are so little affected by heating to temperatures such that a substantial fraction of the material is volatilised and the linear shrinkage may range as high as 5 % (coal K). The most probable explanation would appear to be that the source of the material volatilised is the surface layers of the micelles, which then shrink in together under the mutual action of their surface forces. The abnormally large expansion coefficients recorded on heating (rapidly) to each new temperature level seem to indicate that, prior to evaporation, this material probably forms a mobile adsorbed layer which may also be the seat of chemical changes leading to the formation of distillation products not present in the original material.

Unlike the bituminous coals F and K, the anthracite C was markedly anisotropic. Too much weight should not be attached to the perhaps fortuitous agreement between one of the coefficients and that of graphite normal to the layer planes. It is clear that the regions of continuous covalent network are considerably more extensive in anthracite than in bituminous coal, especially in the direction parallel to the bedding plane.

B.C.U.R.A.,
13 *Grosvenor Gardens,*
*London, S.W.*1.

*GENERAL DISCUSSION**

Dr. W. W. Barkas (*Princes Risborough*) said: The modulus E_A is given by

$$E_A = 100\, F\Sigma/x$$

but by Magg's equation (1) this is identical with

$$E_A = -100\Delta G/x$$

so that this modulus can be derived directly from the free energy decrement and the linear extension, independently of the theory of *surface* free energy changes. In this form attention is drawn to the fact that E_A is actually the ratio of a hydrostatic stress to a directional strain, and it is for this reason that it cannot be identified with Young's modulus where the stress and strain are both directional. The connection between E_A and E_M is through the Poisson's ratios which would require independent measurement.

---

* On two preceding papers.

These observations are not intended to diminish the importance attached to the modulus E_A. The elastic moduli commonly accepted are purely arbitrary choices based on convenience, and any other form of modulus has a right to recognition if its usefulness can be established, but it is an advantage to have the connection between the usual and special moduli clearly set up.

Prof. G.-M. Schwab (*Athens*) (*communicated*) : I want to direct attention to an interesting consequence of the adsorption swelling of charcoal, discovered in 1941 in our laboratory, and, because of war conditions only published in a German periodical.[1] A sample of charcoal grains, enclosed in a constant volume, or put under a constant mechanical pressure, on adsorption of vapours shows an enormous increase of electrical conductivity, increasing with mechanical pressure and proportional to the adsorbed amount. The reason obviously is an increase of contact areas between the swelling grains (microphone effect), and thus swelling is most sensitively indicated by the conductivity. The phenomenon may be of practical use for automatic vapour detection devices.

Mr. F. A. P. Maggs (*London*) (*communicated*) : The derivation of E_A for the simple cases of thin cylinders or laminæ does not involve quite such an arbitrary definition of the modulus as Dr. Barkas suggests. For such models the surface pressure acts in the direction of the resulting extension, and E_A is the quotient of the directed stress divided by the corresponding strain. The non-identity with Young's modulus arises from the fact that whereas the cross-section of a thin cylinder strained in mechanically applied tension will decrease to an extent depending upon Poisson's ratio, there would be a superimposed radial expansion in the case where the increase in length is due to film formation. (In this connection I am asked by Dr. Bangham to correct a mis-statement appearing in a footnote on p. 29 of the Proc. of the B.C.U.R.A. Conference on the Ultra-fine Structure of Coals and Cokes (1943), with which Dr. Barkas's name is incorrectly associated.)

With such a material as charcoal, the measurement even of Young's modulus is not without its difficulties, and that of Poisson's ratio was for this reason not attempted. Had the underlying theory been incorrect there is no reason why E_A and E_M should agree even to an order of magnitude, but with charcoal the agreement is very much better than this.

---

[1] Schwab and Karkalos, *Z. Elektrochem.*, 1941, 47, 345.

GREASE LUBRICATION—THE SWELLING OF SOAPS IN NON-AQUEOUS LIQUIDS.

By E. W. J. Mardles and E. B. Clarke.

Received 2nd August, 1946.

A study of the swelling and solubility behaviour of soaps in liquids other than water is of importance in connection with the preparation of greases and lubricants, and thickened liquids in general, and this work forms part of an investigation on the rheological properties of these materials for aeronautical purposes.

The majority of soaps do not dissolve in organic liquids at room temperature beyond a few mg. per 100 cc., although they show signs of swelling; McBain pointed out, for example,[1] that thorium palmitate and silver palmitate are not soluble in xylene at 25° c., but swell a little and go into solution on heating, setting to jellies on cooling while sodium laurate, potassium palmitate, ferric oleate and barium palmitate swell but do not dissolve in hot xylene. In xylene, McBain found the approximate order of swelling, in increasing amount, for several soaps to be as follows: lithium palmitate, sodium palmitate, potassium oleate, sodium stearate, calcium palmitate, sodium oleate, sodium erucate and chromium palmitate.

In our investigation we have found that the sodium and potassium soaps tried, dried in powder form at 80° for two hours, appear in general, to swell less in organic solvents than the soaps of lithium and of the heavier metals, the stearates less than the palmitates. These differences emphasise that the properties of soaps depend considerably on their method of manufacture and the presence of acids, alkalis, water, etc.; it has been found possible to prepare some of the soaps in various degrees of consistency with differences in the swelling and solubility behaviour in mineral oils. In this work the soaps were supplied by the B D.H., and were dried for about two hours at 80° c.

The behaviour of soaps in liquids with change of temperature and kind of metal present is highly complex, due in part to the mesomorphic changes which vary from soap to soap; Lawrence,[2] has shown that zinc oleate and lead stearate do not exhibit the plastic phase as many soaps do on heating, and according to McBain,[1] aluminium dipalmitate in dry xylene requires a minimum concentration of 12 % before a clear jelly is obtainable, whereas a xylene gel of 0·2 to 0·5 % silver palmitate remains unaltered for weeks.

Greases prepared from lithium stearate (specification D.T.D. 577) have a smaller temperature gradient for consistency than corresponding greases prepared from sodium or calcium stearate, and so are more suitable for low temperature lubrication than, say, a lime base grease which hardens rapidly below 0° c., but why this is so is not clearly understood.

The complex relationship between the rheological and lubricating properties of greases and their fibrous structure, smectic state, etc., is beyond the scope of this paper, but it can be shown that the swelling of the soap which occurs in the hydrocarbons and is in great measure irreversible with lowering of temperature, is of first importance in the problem. The presence of small amounts of water, glycerol, fatty acids, etc., in greases greatly affects the swelling and with a non-polar oil such as a white oil, it is necessary to add water and glycerol, etc., in order to obtain a jelly, with sodium stearate.

Recent X-ray photographic investigations by Bernal, Harkins, and others,[3] have shown that the soap molecules are packed in the smectic state in double molecules with the carboxyl groups contiguous; allotropy occurs by variations in the angle of inclination in the smectic planes. Harkins has measured the distances between the micelles of potassium laurate when polar molecules such as water are allowed to enter the system.

In this paper are presented swelling data obtained with various soaps in organic liquids, alone and in mixtures, over a range of temperature from 20° c. to 120° c. with observations on gelation, dispersion, contact angles, etc. Evidence of the unidimensional melting of the soap molecules has been obtained from the swelling data.

Experimental.—When fine dry sodium stearate powder is shaken with a mineral lubricating oil or other organic liquid, little or no change in the soap is usually visible, but the volume it occupies on long settling is measurably greater than in the dry state, and varies from liquid to liquid, being usually greater with " polar " hydrocarbons and at higher temperatures. On warming above 70° c. the soap, if it does not dissolve, may swell sufficiently to occupy the whole system. The sediment volume of some soaps, such as copper or lead stearate, increases at first with temperature rise then diminishes because the soap tends to dissolve at the higher temperatures. On cooling the soap/liquid systems the sediment volume of the soap does not diminish appreciably, and where partial solution has occurred the solution either forms a jelly or deposits soap in a gelatinous form.

---

[1] McBain and McClatchie, *J. Physic. Chem.*, 1932, 36, 2567.

[2] Lawrence, *Trans. Faraday Soc.*, 1933, 29, 1074; 1938, 34, 660.

[3] Harkins, Matloon and Corrin, *J. Amer. Chem. Soc.*, 1946, 68, 220.

The experimental work consisted in determining the sediment volume in cc. per gram of soap, and the swelling expressed as the sediment volume over and above the dry volume, when 1 gram of powder was shaken with 20 ml. or more of the organic liquid and allowed to settle for at least 24 hours in a tube 10 mm. diameter, under isothermal conditions. The effect of successive temperature rises was observed.

W. Gallay and I. E. Puddington [4] measured the sedimentation volume of powdered sodium oleate in four different mineral oils at various temperatures, and showed that the oil with lowest viscosity index gave the highest sedimentation volume, and that the volume was doubled when the temperature was raised from 25° c. to 100° c. With the same oil they found that sodium oleate occupied a greater volume than the stearate and the effect of adding free fatty acid was to increase the volume and free alkali to decrease it.

Sodium Stearate in Mineral Oils.—Little or no swelling data for soaps in mineral oils have been found in the literature, but the solubility of soaps in a refined white oil (Nujol) has been studied by Lawrence,[2] and Gallay and Puddington,[5] and others, who have shown that dissolution usually occurs above 200° c., and that the results are considerably affected by small amounts of water. Sodium stearate apparently forms a stable monohydrate, but Lawrence prepared anhydrous sodium stearate from sodium and stearic acid in the oil, and found that the solution on cooling did not form a jelly but separation occurred and opalescence developed. With the cooling of 1 % solutions Lawrence found breaks in the neighbourhood of 80° c., indicating some change in the allotropy of the sodium soaps. Using calcium stearate, Gallay and Puddington found that at about 100° c. in various mineral oils there was a partial melting of the soap and a mesomorphic change occurred ; on further heating the soap dissolved.

When the sample of sodium stearate powder used was shaken with a refined white mineral oil (B.P. liquid paraffinum) the sediment volume at 20° c. of a gram was 2·5 cc. By raising the temperature and allowing to stand for several days the sediment volume increased as shown in Fig. 1*a* and it is of interest to note that the relation between swelling and temperature can be represented by two lines cutting at about 70° c. the point of unidimensional melting of the soap molecule.

This genotypical or unidimensional melting-point at 70° c. for sodium stearate has been observed by Thiessen,[6] and Gallay and Puddington from density measurements and differential cooling curves.

Other transitions occur at higher temperatures, and indicate stages in the loosening of the molecular aggregates with a consequential greater entry of the oil molecules into the micelles.

Stearic acid itself melts at about 70° c. when the cohesional forces between the polar bonds and the hydrocarbon chains become completely disrupted. Thiessen has shown by optical methods that the empirical rule of Krafft, who postulated many years ago that changes in the state of dispersion of hydrosols of soap corresponding to the melting of the fatty acid of the soap, is actually borne out in practice and that a form of melting is involved.

---

[4] Gallay and Puddington, *Can. J. Res. B.*, 1944, **22**, 16.

[5] Gallay and Puddington, *ibid.*, 173.

[6] Thiessen *et al.*, *Z. physik. Chem. A.*, 1931, **156**, 435 ; 1933, **165**, 453 ; 1935, **174**, 335 ; 1936, **176**, 397 ; *Angew. Chem.*, 1936, **49**, 640.

In Fig. 1*b* are given results of swelling of soaps in oils, and includes powdered stearic acid in water. Potassium stearate behaved similarly to the sodium salt in B.P. paraffinum, and the effect of prolonged drying of sodium stearate at 120° c. was to shift the curve a few degrees without altering any feature. Using a light mineral oil and paraffin oil, respectively, there was increased swelling over that found with B.P. paraffinum when sodium stearate or sodium palmitate was used, the m.p. of palmitic acid being 62·9° c. The greater swelling capacity of the palmitates compared with the stearates are shown for several hydrocarbons in Fig. 2. Higher swelling was found with xylene than with dekalin, tetralin, white spirit and a light mineral oil, in order of decreasing swelling, for both sodium and sodium palmitate.

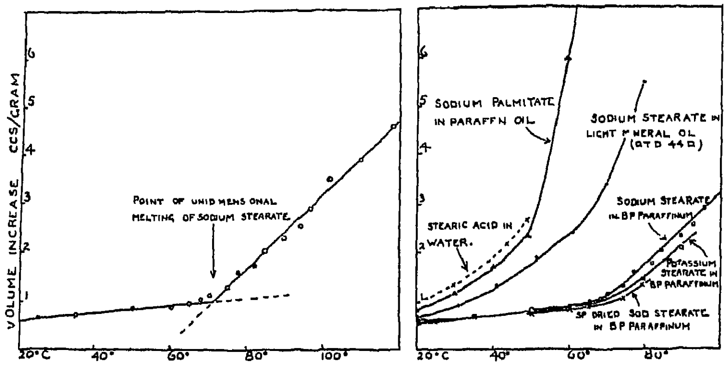
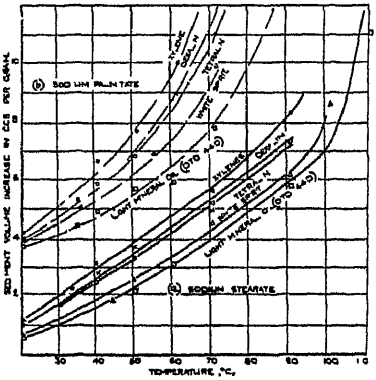


FIG. 1.—The swelling of soaps in mineral oils at various temperatures.



THE SWELLING OF (a) SODIUM STEARATE
(b) SODIUM PALMITATE, IN VARIOUS HYDROCARBONS
AT TEMPERATURES TO 110°C.

FIG. 2.

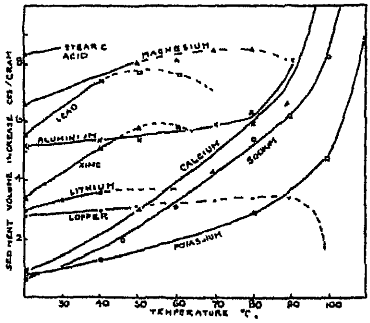


FIG. 3.—The swelling of stearates in white spirit (B.S.S. 245) at temperatures to 110° c.

The Swelling of Stearates in White Spirit.—Several metallic stearates were examined in white spirit at temperatures to 110° c. Of the soaps examined at 20° c., potassium, sodium and ammonium stearate exhibited the least swelling; slightly increased swelling occurred with calcium stearate, and in order of increasing swelling then followed copper with a sediment volume increase over the dry volume of 2·8 cc./g., lithium 2·9 cc./g., zinc 3·5 cc./g., aluminium 5·1 cc./g., lead 5·6 cc./g. and magnesium 6·8 cc./g., the stearates of Cu, Zn, Pb, and Mg. began to dissolve in the neighbourhood of 50° c. with a consequential decrease in the sediment volume with rise of temperature (Fig. 3).

TABLE I

The influence of small additions on the swelling of stearate of sodium, copper, aluminium and zinc respectively in (*a*) light mineral oil (D.T.D. 44D), and (*b*) dekalin. Sediment volume increase is in cc. per gram of soap above the dry volume.

	2 days at 20° c.	+ 0·1 % (based on soap) Oleic Acid.	+ 0·1 Oleic Acid + 0·3 % Triethanolamine.
(*a*) *D.T.D. 44D Oil*			
Sodium stearate .	0·7	6·2	9·7
Copper „ .	2·5	2·8	2·8
Aluminium „ .	4·9	7·4	12·5
Zinc „ .	6·9	11·6	12·4
(*b*) *Dekalin*			
Sodium stearate .	1·1	2·6	5·4
Copper „ .	2·3	2·9	3·0
Aluminium „ .	4·5	4·5	9·5
Zinc „ .	5·5	11·2	12·5

Effect of Small Additions.—The large effects of small additions on the swelling of stearates is shown in Table I, with four soaps in (*a*) light mineral oil (D.T.D. 44D); and (*b*) dekalin, and in Table II for water in benzene systems, at 20° c.

In general, the effect of small amounts of water, glycerol, oleic acid, triethanolamine, etc., which cause increased swelling, on the consistency of the grease was to soften it, reducing the rigidity. Glycerol which seems to behave as a plasticising agent for the soap reduces the jelly strength, but raises slightly the melting point of the jelly formed with sodium stearate in a white mineral oil. Gallay and Puddington [7] found that with sodium stearate in a mineral oil the addition of 10 % (based on the soap) of glycerol produced an increase in viscosity at about 95° c., and on cooling from this temperature short fibres were noted. With phenol a very long fibred grease was obtained. Glycols and water were all found to show increases in viscosity and produced fibres on cooling.

TABLE II.

The swelling of sodium stearate in benzene with the addition of small amounts of water, at 20° c.

Sed. vol. cc./g., sod. stearate in dry benzene	2·3
+ 0·05 cc. water in 100 cc. dry benzene	2·9
+ 0·1	4·2
+ 0·25	5·8
+ 0·5	11·1

Swelling of Soaps in Single Liquids.—The swelling effect with various organic liquids could be followed using the sedimentation method; with viscous liquids such as glycerol it was essential to allow ample time for settling. Some of the data obtained are given in Table III; angles of contact of the liquids on a film of sodium stearate are included as adhesion tensions (Freundlich) namely $\gamma \cos \theta$ where γ is surface tension and

---

[7] Gallay and Puddington, *Can. J. Res. B.*, 1943, **21**, 211.

θ the contact angle. A film of soap on glass was obtained by coating the warm glass surface with a hot dilute solution of the soap in a mixture of

TABLE III.

The swelling (sediment volume increase, cc./g., above dry volume), of (i) sodium stearate, and (ii) lithium stearate in organic liquids; $\gamma \cos \theta$ the adhesion tension (Freundlich) of liquid to soap is included.

	(i) Sodium Stearate.		(ii) Lithium Stearate.	
	Swelling.	Adhesion Tension.	Swelling.	Adhesion Tension.
Petroleum ether	1·0	—	—	—
White spirit	0·9	22·6	3·1	24·9
Light mineral oil (D.T.D. 44D)	0·9	19·8	3·8	—
White mineral oil (B.P. paraffinum)	0·7	19·2	2·8	20·5
Aero engine oil	0·8	19·5	—	—
Chloroform	0·5	—	—	—
Carbon tetrachloride	0·7	—	—	—
Benzene	1·3	27·8	3·8	28·4
Toluene	1·2	27·6	3·5	27·2
Tetralin	1·3	29·2	—	28·9
Ethyl alcohol (abs.)	0·9	22·1	3·0	19·3
n-Prophyl alcohol	1·5	23·7	2·9	20·0
Cyclohexanol	2·3	20·2	3·4	—
Glycerol	2·2	15·1	Does	not wet
Ethyl laurate	1·2	—	2·3	—
Butyl stearate	1·2	—	—	—
Acetone	0·9	—	0·2	—
Water	above 100	—	Does	not wet

chloroform and alcohol and drying off in an incubator at 80° c.; the angles were measured using a low powered microscope and a goniometer.

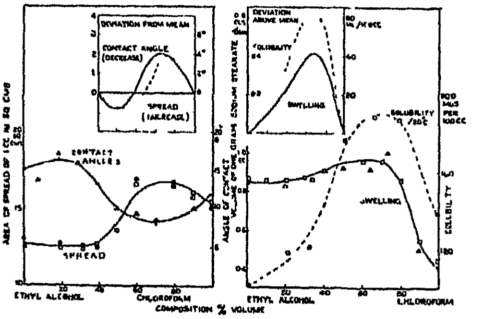


FIG. 4.—Contact angles and area of spread of ethyl alcohol-chloroform mixtures on sodium stearate; also swelling and solubility of sodium stearate in the mixtures.

With sodium stearate, swelling and adhesion tensions are higher with the aromatic hydrocarbons than with the paraffin hydrocarbons.

Swelling Temperature Relation.—The relationship between swelling and temperature is best seen with a "non-polar" liquid such as the refined white mineral oil, used for the experiments, the results of which are shown in Fig. 1. Sodium stearate in the white mineral oil does not begin to dissolve apparently until a temperature above 120° c. is reached; with alcohols and the aromatic hydrocarbons the soaps begin to dissolve on warming above 50°, so the increased swelling is offset by the loss of soap dispersed, and the curves relating the sediment volume and temperature often show a decreased rate of increase of swelling with temperature rise.

Swelling in Mixed Liquids.—It is to be expected as with other colloids in mixed liquids,[8] that the swelling of soaps is much greater in the mixtures than in the single liquids; this behaviour was found in several mixtures, namely, benzene and alcohol, alcohol and water, alcohol and chloroform, oleic acid and white spirit, alcohol and glycerol, and others.

S. R. Palit[9] has found that sodium stearate in mixed ethylene glycol and butyl alcohol can dissolve in amounts up to 10 % at room temperature, although the single liquids are indifferent solvents.

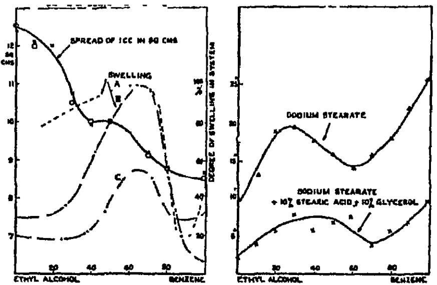


FIG. 5.—Area of spread and contact angles of ethyl alcohol-benzene mixtures on sodium stearate; also swelling after 3 months, and after 3 days.

Concomitant with the increased swelling and dispersion is a decrease in contact angles and increase in the area of the pool of liquid formed on the flat soap surface, using the same volume of liquid, thus, as shown in Fig. 4, with sodium stearate in mixtures of ethyl alcohol and chloroform maximum swelling occurs with the mixture of maximum solvent power which also has the lowest contact angle and maximum area of the pool for 1 cc. The same behaviour is found with ethyl alcohol and benzene, Fig. 5, and ethyl alcohol and water, Fig. 6.

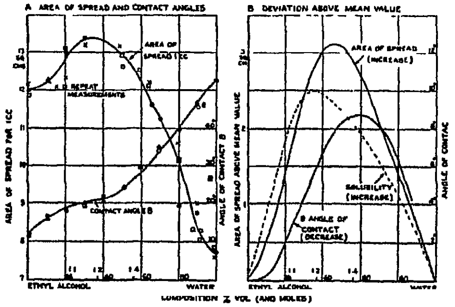


FIG. 6.—Area of spread and contact angles of ethyl alcohol-water mixtures on sodium stearate; also solubility of sodium stearate in the mixtures.

Consistency of Soap Thickened Liquids.—A simple method of determining the consistency of the various soap systems at temperatures to − 80° c. has been worked out, using a thin plate passing through the grease edge on.[10] The relation between the swelling of various metal soaps, fibre length, etc., and the rheological properties of the grease is being worked out.

Summary.

The paper describes an investigation on the swelling of various soaps in organic liquids at temperatures to 120° c., using a sedimentation of powder method. With the swelling data has been included contact angles and adhesion tensions (Freundlich).

The effect of small additions of water, oleic acid, glycerol, etc., to the soap systems has been studied, also the behaviour of soaps in mixed liquids. The relationship between the swelling of a soap and the rheological properties of its grease at temperatures to − 80° c. are being worked out.

---

[8] Mardles, *J. Chem. Soc.*, 1925, **127**, 2940.

[9] Palit, *J. Indian Chem. Soc.*, 1942, **19**, 271.

[10] Mardles, *Nature*, 1946, **158**, 199.

GENERAL DISCUSSION[*]

Dr. J. Th. G. Overbeek (*Eindhoven*) said: In connection with the experiments of Drs. Mardles and Clarke on the influence of small additions of water on the swelling of sodium stearate in benzene, I want to draw the attention to the work of Kruyt and van Selms,[1] who investigated the rheological properties of suspensions of quartz and starch in nonpolar liquids. They found that the suspensions were stable (in the colloid chemical sense of the word) when the system was absolutely dry but unstable when small quantities of water were present. Their explanation was, that the water formed thin layers around the (hydrophilic) particles and that the water films would coalesce, when two particles touched each other thereby diminishing the free surface energy of the system.

Now unstable suspensions show a larger sediment volume than stable ones because the coagulation of the particles prevents the most advantageous packing. As in the experiments of Mardles and Clarke swelling has been measured by determining the sediment volume, it is quite possible that the enhanced swelling of sodium stearate in benzene in the presence of water is wholly or partly due to the same effect found by Kruyt and van Selms and not to a proper swelling of the soap.

Mr. V. R. Gray (*Manchester*) said: The effect of small amounts of added substances on the swelling and solution of metallic soaps in hydrocarbons is generally termed peptisation. It is an entirely different effect from peptisation of colloids in aqueous media, however, and it obviously does not involve ionic dissociation. Its explanation can be found by observing the type of substances which behave as powerful peptising agents. The strongest are substances such as amines and water, that is, substances which easily enter into co-ordination complexes; then come alcohols, phenols, and finally fatty acids. That is, increase of acid strength, or decrease of ease of donation of electrons, decreases the peptising ability. It therefore appears that a peptiser acts by preferentially co-ordinating itself on to the soap molecule, thus breaking the mass into small units allowing swelling and suspension or aiding solution and decrease of viscosity. This explanation is particularly applicable to aluminium soap solutions for which a co-ordinated structure is proposed but seems to hold generally for soaps in non-aqueous media.

Dr. J. B. Matthews (*Chester*) (*communicated*): In considering the effects produced by the addition of small amounts of water and other agents on the swelling behaviour of soaps in organic liquids and the stability of the resulting gels, it is important to consider these quantities in terms of molecular proportions rather than proportions by weight. Thus, it is well-known that to obtain stable calcium soap greases it is necessary to incorporate about 0·5-1·0 % by weight of water in the grease. This corresponds roughly to one molecule of water per molecule of soap, and according to the accepted structure of the soap micelle there is sufficient water present to form a layer of one molecule thickness between the opposing layers of the polar heads of the soap molecules. If this is the correct picture, it would appear that the presence of the intercalated monomolecular water layer favours lateral end-to-end adherence of the micelles, producing the long chains which are responsible for the gel structure, whereas, in its absence perpendicular adherence results and large crystals of soap separate from the oil.

---

[*] On preceding paper.

[1] Kruyt and van Selms, *Rec. trav. Chim.*, 1943, **62**, 398, 407, 415.

Contents

PAGE

General Introduction. By J. D. Bernal 1

Diffusion and Swelling of High Polymers. Part I: The Swelling and Solution of a High Polymer Solid considered as a Diffusion Process. By G. S. Hartley 6

Diffusion and Swelling of High Polymers. Part II: The Orientation of Polymer Molecules which accompanies Uni-directional Diffusion. By Conmar Robinson 12

The Reactions Occurring in the Gelatinisation of the Nitrocelluloses. By Miss T. Petitpas and M. Mathieu 17

General Discussion.—Dr. R. J. B. Marsden, Dr. J. J. Hermans, Dr. G. S. Hartley, Mr. J. Crank, Dr. J. R. Scott, Dr. P. H. Hermans, Dr. C. Robinson 29

The Interaction between Rubber and Liquids. Part X: Some New Experimental Tests of a Statistical Thermodynamic Theory of Rubber-liquid Systems. By G. Gee 33

General Discussion.—Dr. R. F. Tuckett, Prof. E. A. Guggenheim, Dr. M. Magat, Dr. P. M. Doty, Dr. J. J. Hermans, Dr. J. R. Scott, Dr. G. Gee, Dr. G. Salomon, Prof. E. C. Baughan . 44

The Effect of Solvent-Type and Temperature on the Specific Viscosities of Polystyrene Solutions. By T. Alfrey, J. D. Justice and S. J Nelson 50

Particle Swelling and the Viscosity of Suspensions. By F. Eirich and J. Sverak 57

General Discussion.—Dr. J. J. Hermans, Mr. V. R. Gray, Dr. E. M. Frith 65

The Determination of Macromolecular Configurations in Dilute Solution by Light Scattering. By P. M. Doty, W. A. Affens and B. H. Zimm 66

General Discussion.—Prof. P. Putzeys, Dr. G. Gee, Dr. Magat, Dr. P. M. Doty, Dr. F. Eirich, Mr. F. W. Peaker . . . 78

Mutual Solubility of Liquids. By A. E. Van Arkel . . . 81

Absorption of Water by Polymers: Analysis in Terms of a Simple Model. By A. J. Hailwood and S. Horrobin 84

General Discussion.—Dr. G. Salomon, Dr. G. Gee, Dr. M. Magat, Dr. P. H. Hermans, Dr. W. W. Barkas, Dr. R. M. Barrer, Dr. Vermaas, Prof. A. R. Ubbelohde, Dr. L. R. G. Treloar, Mr. A. J. Hailwood, Mr. Horrobin 93

On the Interaction of Colloidal Particles. Part I: Particular Application to Parallel Plates. By S. Levine 102

Long Distance Forces Acting between Colloidal Particles. By E. J. W. Verwey and J. Th. G. Overbeek 117

General Discussion.—Dr. M. F. Perutz, Dr. F. Eirich, Mr. V. R. Gray, Dr. D. D. Eley, Dr. J. Th. G. Overbeek, Dr. S. Levine, Prof. J. D. Bernal 123

A Method for Measuring the Transverse Swelling of Fibres. By J. Preston 131

Sorption, Swelling and Elastic Constants of the Cell Wall Material in Wood. By W. W. Barkas 137

Swelling Stresses in Anisotropic Cylinders. By F. L. Warburton . 151

The Structure and the Mechanism of Deformation in Cellulose gels. I: General Picture of a Molecular Network. By P. H. Hermans and D. Vermaas 155

PAGE
The Structure and the Mechanism of Deformation in Cellulose gels. II: Application of the Theory to Short Chains. By J. J. Hermans 160
General Discussion.—Dr. R. J. B. Marsden, Mr. Warburton, Dr. A. R. Urquhart, Mr. F. A. P. Maggs, J. St. L. Philpot, Dr. W. W. Barkas 163
On the Structure of Transparent Water and Oil Dispersions (Solubilised Oils). By J. H. Schulman and T. S. McRoberts 165
Critical Effect of Temperature on the Absorption of Water by Solutions of Ethanolamine Oleate in Benzene. By R. C Pink . 170
Studies of Aluminum Soaps. VII: Aluminum Soaps in Hydrocarbons. The Gels and Jellies and Transformations between them. By J. W. McBain, K. J. Mysels and G. H. Smith . . 173
Swelling and Molecular Organisation in Colloidal Electrolytes. By D. G. Dervichian 180
The Composition and Swelling Properties of Hæmoglobin Crystals. By M. F. Perutz 187
General Discussion.—Dr. G. S. Hartley, Mr. V. R. Gray, Prof. J. D. Bernal 196
Randomness in the Structures of Kaolinitic Clay Minerals. By G. W. Brindley and K. Robinson 198
On the Hydration of Montmorillonite. By J. Mering 205
Ionic Forces in Thick Films of Liquid between Charged Surfaces. By R. K. Schofield 219
General Discussion.—Dr. J. Th. G. Overbeek, Dr. R. K. Schofield, Mr. D. M. G. MacEwan, Dr. G. W. Brindley, Dr. M. Mathieu 226
The Swelling of Protein Fibres in Organic Solvents. By Miss D. Jordan-Lloyd, Miss M. Dempsey and Miss M. Garrod . . 228
Relation of Swelling to Shrinage Temperature of Collagen. E. R. Theis 244
General Discussion.—Dr. J. L. Stoves, Dr. G. Salomon, Mr. L. Peters, Prof. J. D. Bernal 249
Swelling and Cytoplasmic Vacuolation. By E. Fauré-Fremiet . 253
The Swelling of Starch Grains. By N. P. Badenhuizen . . 255
The Swelling and Shrinking of Protoplasm. By W. Seifriz . . 259
The Swelling and Shrinking of Mammalian Red Cells: Departures from Ideal Osmotic Behaviour. By E. Ponder 266
General Discussion.—Mr. N. Goldenberg, Dr. W. W. Barkas, Mr. L. Peters, Dr. R. K. Schofield, Prof. A. R. Ubbelohde, Dr. D. H. Bangham, Dr. M. F. Perutz, Prof. J. D. Bernal . . . 272
The Swelling of Reinforced Plastics. By F. T. Barwell and K. W. Pepper 275
General Discussion.—Dr. W. W. Barkas, Mr. F. I. G. Rawlins, Dr. K. W. Pepper 282
The Adsorption-Swelling of several Carbonaceous Solids. By F. A. P. Maggs 284
The Thermal Expansion of Coals and Carbonised Coals. By D. H. Bangham and Miss R. E. Franklin 289
General Discussion.—Dr. W. W. Barkas, Prof. G.-M. Schwab, Mr. F. A. P. Maggs 294
Grease Lubrication—the Swelling of Soaps in Non-Aqueous Liquids. By E. W. J. Mardles and E. B. Clarke 295
General Discussion.—Dr. J. Th. G. Overbeek, Dr. V. R. Gray, Dr. J. B. Matthews 302

Transactions of the Faraday Society

INSTRUCTIONS TO AUTHORS

Authors submitting papers for publication are requested to note the following particulars :—

i. All communications made to the Society are the property of the Society. When papers have been accepted for publication, authors are not at liberty, save by permission of the Council, to publish them elsewhere. Such permission will not be unreasonably withheld. Communications which have appeared in any other Journal may not usually be published in the Transactions.

ii. (*a*) **Papers should be as brief as is compatible with clarity,** and no experimental details should be included which are not necessary for the understanding of the paper or for verifying any conclusions that may be drawn. Paragraphs should not be unnecessarily sub-divided.

(*b*) Tables should be reduced to a minimum and, where possible, **results should be shown by means of curves rather than tables.** Only in exceptional cases may tables be published as well as graphical representations of their contents.

(*c*) Each paper should include a brief summary indicating in general terms its purpose and conclusions.

iii. (*a*) Papers must be typewritten, with double spacing, on one side only of the paper, with a good margin on the left side.

(*b*) **References** should be typed in numerical order (and in the following sequence: journal; date; volume number; page) at the end of the paper; the necessary reference numbers only should appear in the text.

(*c*) Symbols, formulæ and equations should be written or typewritten with great care. The symbols recommended in the Report of the Joint committee of the Chemical Society, the Faraday Society and the Physical Society should be employed.

iv. (*a*) **Line drawings** may in the first instance be submitted in sketch form. For publication, however, they must be made with Indian ink on white paper or preferably tracing cloth. Freehand must not be employed and all lines must be firmly and evenly drawn. Drawings should, generally, be capable of reduction to about one-half scale, depending on their nature and complexity; all drawings relating to one paper will be reduced to the same extent. In the case of curves sufficient ordinates and abscissæ to facilitate reading the curves must be drawn in ink and reliance must not be placed on the printed lines of scale paper.

(*b*) Where possible, lettering should be inside the diagram. Lettering should be put in lightly with lead pencil.

v. **Rapidity of publication** is facilitated if papers are submitted in duplicate with duplicate photographs or tracings of diagrams, so that each paper with its drawings and photographs will fold into a small envelope.

vi. Authors must retain copies of their communications.

Correction of Proofs.—The address to which proofs are to be sent should be written on every paper. Authors resident overseas may name agents in Britain to whom may be referred matters concerning their papers, including the correction of proofs, in order that delay in publication may be avoided.

Two galley proofs are provided, one of which duly corrected should be returned to the Editor. Authors are required to submit their communications in such form that corrections other than of printer's errors will be unnecessary. **Authors making other corrections will be required to bear the cost thereof.**

Reprints.—Twenty-five reprints (without covers) are presented to each author **who is a member of the Society,** up to a maximum of fifty reprints in all.